

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 16, 2002, 23:48:11 ; Search time 46 Seconds
(without alignments)
4170.430 Million cell updates/sec

Title: US-09-658-824-808

Perfect score: 781

Sequence: 1 gcggcgagctgtgagccgg.....ctttcccccaaaaaaaaaa 781

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	14.1	539	1	US-08-370-648-18
2	110	14.1	539	2	US-08-531-662B-18
3	110	14.1	539	3	US-08-669-161A-18
4	110	14.1	539	3	US-08-602-039-18
5	110	14.1	540	1	US-08-370-648-16
6	110	14.1	540	3	US-08-531-662B-16
7	110	14.1	540	3	US-08-669-161A-16
8	110	14.1	540	3	US-08-602-039-16
9	108.4	13.9	532	2	US-08-370-648-17
10	108.4	13.9	532	2	US-08-531-662B-17
11	108.4	13.9	532	3	US-08-669-161A-17
12	108.4	13.9	532	3	US-08-602-039-17
13	108.4	13.9	538	1	US-08-370-648-14
14	108.4	13.9	538	2	US-08-531-662B-14
15	108.4	13.9	538	3	US-08-669-161A-14
16	108.4	13.9	538	3	US-08-602-039-14
17	108.4	13.9	560	1	US-08-370-648-15
18	108.4	13.9	560	2	US-08-531-662B-15
19	108.4	13.9	560	3	US-08-669-161A-15
20	108.4	13.9	560	3	US-08-602-039-15
21	89.6	11.5	515	4	US-09-439-313-472
22	75.4	9.7	11461	3	US-08-669-161A-29
23	68	8.7	7218	1	US-08-232-463-14
24	58.4	7.5	646	1	US-08-370-648-1
25	58.4	7.5	646	2	US-08-531-662B-1
26	58.4	7.5	646	3	US-08-669-161A-1
27	58.4	7.5	646	3	US-08-602-039-1

28	58.4	7.5	648	1	US-08-250-162A-1	Sequence 1, Appli
29	38	4.9	289	4	US-09-007-005-17	Sequence 17, Appl
30	38	4.9	289	4	US-09-244-796-17	Sequence 17, Appl
c 31	36.8	4.7	5288	2	US-08-540-406-18	Sequence 18, Appl
c 32	36.8	4.7	5288	3	US-08-656-055-18	Sequence 18, Appl
c 33	36.8	4.7	5288	4	US-08-954-668-18	Sequence 18, Appl
c 34	36.8	4.7	5288	5	PCT-US95-13233-18	Sequence 18, Appl
c 35	36	4.6	48974	4	US-08-920-422-17	Sequence 17, Appl
c 36	34	4.4	43280	2	US-08-804-227C-1	Sequence 1, Appl
37	33	4.2	3383	1	US-07-707-367-1	Sequence 1, Appl
38	32.8	4.2	1693	2	US-08-487-113D-118	Sequence 118, App
39	32.8	4.2	1693	2	US-08-720-420A-118	Sequence 118, App
40	32.4	4.1	3251	4	US-09-085-199B-6	Sequence 6, Appli
41	32	4.1	954	3	US-09-418-641-3	Sequence 3, Appli
c 42	32	4.1	1437	1	US-08-252-966B-14	Sequence 14, Appl
43	32	4.1	4695	2	US-08-231-193A-57	Sequence 57, Appl
44	32	4.1	4695	2	US-08-486-273A-57	Sequence 57, Appl
45	32	4.1	4695	3	US-08-940-086A-57	Sequence 57, Appl

ALIGNMENTS

RESULT 1

US-08-370-648-18

; Sequence 18, Application US/08370648

; Patent No. 5648226

; GENERAL INFORMATION:

; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;

; APPLICANT: Boon-Falleur, Thierry;

; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid

; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen, and Uses Thereof

; NUMBER OF SEQUENCES: 18

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felfe & Lynch

; STREET: 805 Third Avenue

; CITY: New York City

; STATE: New York

; COUNTRY: USA

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; FILING DATE: 10-JANUARY-1995

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/250,162

; FILING DATE: 27-MAY-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/096,039

; FILING DATE: 22-JULY-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Hanson, No. 5648226man D.

; REGISTRATION NUMBER: 30,946

; REFERENCE/DOCKET NUMBER: LUD 5323.2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 688-9200

; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 18:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 539 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-370-648-18

Query Match 14.18; Score 110; DB 1; Length 539;

Best Local Similarity 63.9%; Pred. No. 3.6e-25;


```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/370,648
; FILING DATE: 10-JANUARY-1995
; APPLICATION NUMBER: 08/250,162
; FILING DATE: 27-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-JULY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6069001man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5323.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 539 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-602-039-18

Query Match 14.1%; Score 110; DB 3; Length 539;
Best Local Similarity 63.9%; Pred. No. 3.6e-25;
Matches 216; Conservative 0; Mismatches 115; Indels 7; Gaps 3;

QY 447 CCCCAAAAGAACAGACAGCAGCTGAAGTCGGGATCCTACACCTGGCAGCACACAGAA 506
   || || || || || || || || || || || || || || || || || || || ||
Db 202 CCTGAAGAAGGGAACAGCAACTCAACGTCAGAGTCTCGAGCTGCTCAGGAG---GGA 258
   || || || || || || || || || || || || || || || || || || || ||
QY 507 GAAGATCAGGATACAGCTAGATCCCAAGTCCCGACATGGAAGGTGATCTGCAAGAGCTG 566
   || || || || || || || || || || || || || || || || || || || ||
Db 259 GAGGATCAGGAGCACTCGCAGTCAAGGCCGAGCCTGAAGCTGTATAGCCAGGAACAG 318
   || || || || || || || || || || || || || || || || || || || ||
QY 567 CATCAGTCAACACCGGGGATAAATCTGGATTTGGGTTCCGGCTCAAGGTGAAGATAAT 626
   || || || || || || || || || || || || || || || || || || || ||
Db 319 GGTCAACCCACAGACTGGGTGTGAGTGTGAAGAT--GGTCCTGATGGCAGGAGGTGGACCC 377
   || || || || || || || || || || || || || || || || || || || ||
QY 627 ACCTAAGAGGAACACTGTAAATGCCAAGACGAGTGAAGAGCAACACACAGTTTAAAT 686
   || || || || || || || || || || || || || || || || || || || ||
Db 378 GCCAATATCCAGGAGGTGAAACGCTGAAGAAGGTGAAGAAGCAATCACAGTGTATAAA 437
   || || || || || || || || || || || || || || || || || || || ||
QY 687 GAAGACAAGCTGAACACCAACGCAAGCTGGTTTATATATAGATATTG---ACTTAAACTAT 743
   || || || || || || || || || || || || || || || || || || || ||
Db 438 GAAGACAGCTTGAATATGATCGACGCTCTCTCTATGTTGAAATTTGTTTCATTAAATTC 497
   || || || || || || || || || || || || || || || || || || || ||
QY 744 CTCAATAAAGTTTTGCAGCTTTTCACCAAAAAAANAANA 781
   || || || || || || || || || || || || || || || || || || || ||
Db 498 CCAATAAAGCTTTACAGCCTTCTGCAAAAAAANAANAANA 535
   || || || || || || || || || || || || || || || || || || || ||

RESULT 5
US-08-370-648-16
; Sequence 16, Application US/08370648
; Patent No. 5648226
; GENERAL INFORMATION:
; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
; APPLICANT: Boon-Falleur, Thierry;
; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: tm ps/2

```

OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/370.648
FILING DATE: 10-JANUARY-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/250.162
FILING DATE: 27-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/096.039
FILING DATE: 22-JULY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5648226man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5323.2
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 540 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-370-648-16

Query Match 14.1%; Score 110; DB 1; Length 540;
Best Local Similarity 63.9%; Pred. No. 3.6e-25;
Matches 216; Conservative 0; Mismatches 115; Indels 7; Gaps 3;
QY 447 CCCCCAAAAGAACACAGAGCTGAAAGTCGGGATCCTACACCTGGCGCAGACACAA 506
DB 203 CCGTGAAGAGGGAACACCACTCAACGTCAGGATCTCGAGCTGCTCAGGAG---GGA 259
QY 507 GAAGATCAGGATACAGTCAGATCCAGTCGCGGACATGGAAGTGATCTGCAAGAGCTG 566
DB 260 GAGGATGAGGAGGAGTCGTCAGGTCAGGCGGCAAGCTGAAGCTGATAGCCAGGAACAG 319
QY 567 CATCAGTCAAAACACCGGGGATAATCTGGATTTCGGCTTCCGGCTCAAGGTGAAGATAAT 626
DB 320 GGTCACCCACAGACTGGGTGTGAGTGAAGAT---GGTCCTGATGGCGAGGATGGACCC 378
QY 627 ACCTAAGAGGAACTGTAAAATGCCAAGCAGCTGGAAGAGCAACCAAGTTTAAAT 686
DB 379 GCCAAATCCAGAGGAGGTGAAACGCCCTGAAGAAGGTGAAAGCAATCAGAGTGTAAAA 438
QY 687 GAAGACAAGCTGAACACACCAAGCTGTTTATATTAGATATTG---ACTTAAACTAT 743
DB 439 GAAGGCA-GTTGAATGATGCGGCTCTCTATGTTGGAAATTTGTCATTAAATTTCT 498
QY 744 CTCAATAAAGTTTTGCAGCTTTCACCAAAAAA 781
DB 499 CCCAATAAAGCTTTACAGCCTTCTGCAAAAAA 536

RESULT 6
US-08-531-662B-16
Sequence 16, Application US/08531662B
Patent No. 5858689
GENERAL INFORMATION:
APPLICANT: Van der Bruggen, Pierre;
APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
APPLICANT: Boon-Falleur, Thierry
TITLE OF INVENTION: Isolated, Truncated Nucleic
TITLE OF INVENTION: Acid Molecules Which Code For GAGE Tumor
TITLE OF INVENTION: Rejection Antigen, The Tumor Rejection Antigen,
TITLE OF INVENTION: and Uses Thereof
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felie & Lynch
STREET: 805 Third Avenue
CITY: New York City

STATE: New York
COUNTRY: USA
ZIP: 10024
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/531.662B
FILING DATE: 21-September-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/370.648
FILING DATE: 10-January-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/250.162
FILING DATE: 27-May-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/096.039
FILING DATE: 22-July-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5858689man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5323.3
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 540 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-531-662B-16

Query Match 14.1%; Score 110; DB 2; Length 540;
Best Local Similarity 63.9%; Pred. No. 3.6e-25;
Matches 216; Conservative 0; Mismatches 115; Indels 7; Gaps 3;
QY 447 CCCCCAAAAGAACACAGAGCTGAAAGTCGGGATCCTACACCTGGCGCAGACACAA 506
DB 203 CCGTGAAGAGGGAACACCACTCAACGTCAGGATCTCGAGCTGCTCAGGAG---GGA 259
QY 507 GAAGATCAGGATACAGTCAGATCCAGTCGCGGACATGGAAGTGATCTGCAAGAGCTG 566
DB 260 GAGGATGAGGAGGATCTGTCAGGTCAGGCGGCAAGCTGAAGCTGATAGCCAGGAACAG 319
QY 567 CATCAGTCAAAACACCGGGGATAATCTGGATTTCGGCTTCCGGCTCAAGGTGAAGATAAT 626
DB 320 GGTCACCCACAGACTGGGTGTGAGTGAAGAT---GGTCCTGATGGCGAGGATGGACCC 378
QY 627 ACCTAAGAGGAACTGTAAAATGCCAAGCAGCTGGAAGAGCAACCAAGTTTAAAT 686
DB 379 GCCAAATCCAGAGGAGGTGAAACGCCCTGAAGAAGGTGAAAGCAATCAGAGTGTAAAA 438
QY 687 GAAGACAAGCTGAACACCAAGCTGTTTATATTAGATATTG---ACTTAAACTAT 743
DB 439 GAAGGCACTGTTGAATGATGCGGCTCTCTATGTTGGAAATTTGTCATTAAATTTCT 498
QY 744 CTCAATAAAGTTTTGCAGCTTTCACCAAAAAA 781
DB 499 CCCAATAAAGCTTTACAGCCTTCTGCAAAAAA 536

RESULT 7
US-08-669-161A-16
Sequence 16, Application US/08669161A
Patent No. 6013481
GENERAL INFORMATION:
APPLICANT: DeBacker, Olivier; Van den Eynde,
APPLICANT: Benoit; Boon-Falleur, Thierry
TITLE OF INVENTION: Isolated, Truncated Nucleic Acid

;; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
;; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
;; NUMBER OF SEQUENCES: 29
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Felfe & Lynch
;; STREET: 805 Third Avenue
;; CITY: New York City
;; STATE: New York
;; COUNTRY: USA
;; ZIP: 10022
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
;; COMPUTER: IBM PS/2
;; OPERATING SYSTEM: PC-DOS
;; SOFTWARE: Wordperfect
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/669,161A
;; FILING DATE: 24-June-1996
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/531,662
;; FILING DATE: 21-September-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/370,648
;; FILING DATE: 10-January-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/250,162
;; FILING DATE: 27-May-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/096,039
;; FILING DATE: 22-July-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hanson, No. 6013481man D.
;; REGISTRATION NUMBER: 30,946
;; REFERENCE/DOCKET NUMBER: LUD 5443
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 688-9200
;; TELEFAX: (212) 838-3884
;; INFORMATION FOR SEQ ID NO: 16:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 540 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; US-08-669-161A-16

Query Match 14.1%; Score 110; DB 3; Length 540;
Best Local Similarity 63.9%; Pred. No. 3.6e-25;
Matches 216; Conservative 0; Mismatches 115; Indels 7; Gaps 3;

QY 447 CCCCCAAAAGAACACAGCAGCTGAAAGTCGGATCCTACACCTGGCGCAGCAGACAGAA 506
DB 203 CCTGAAGAGGGGAACACAGCACTCAAGCTCAGGATCCTGCAGCTGCTCAGGAG---GGA 259
QY 507 GAAGATCAGGATACAGCTGAGATCCCAGTCCGACATGGAAGGTGATCTCCAGAGAGCTG 566
DB 260 GAGGATGAGGAGGATCTGCAAGTCAAGGGCCGAAGCCTGAAGCTGATAGCAGGAACAG 319
QY 567 CATCAGTCAAAACACCGGGGATAATCTGGATTTGGTTCCGCGCTCAAGGTGAAGATAAT 626
DB 320 GGTACCCACAGACTGGGTGTGAGTGTGAAGAT---GGTCTGATGGCAGGAGATGGACCC 378
QY 627 ACCTTAAGAGGAACACTGTAAATGCCAGAGCAGGTGAAGAGCAACCAACCAAGTTTAAT 686
DB 379 GCCAAATCCAGAGGAGGTGAAACGCCCTGAAGAGGTTGAAAGCAATCACAGTCTTAAAA 438
QY 687 GAAGCAGAGCTGAACACCAACCAAGCTGGTTTATATAGATATTG---ACTTTAACTAT 743
DB 439 GAAGCAGGTTGAAATGATGAGGCTGCTCTATATGTTGGAATTTGTTCAATTAATAATCT 498
QY 744 CTCATTAAGTTTTCAGCTTTTACCAAAAAA 781
DB 499 CCAATAAGCTTTACAGCCTTCTGCAAAAAA 536

RESULT 8

US-08-602-039-16
;; Sequence 16, Application US/08602039
;; Patent No. 6069001
;; GENERAL INFORMATION:
;; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
;; APPLICANT: Boon-Falleur, Thierry;
;; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
;; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
;; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
;; NUMBER OF SEQUENCES: 18
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Felfe & Lynch
;; STREET: 805 Third Avenue
;; CITY: New York City
;; STATE: New York
;; COUNTRY: USA
;; ZIP: 10022
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
;; COMPUTER: IBM PS/2
;; OPERATING SYSTEM: PC-DOS
;; SOFTWARE: Wordperfect
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/602,039
;; FILING DATE: 15-FEB-1996
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/370,648
;; FILING DATE: 10-JANUARY-1995
;; APPLICATION NUMBER: 08/250,162
;; FILING DATE: 27-MAY-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/096,039
;; FILING DATE: 22-JULY-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hanson, No. 6069001man D.
;; REGISTRATION NUMBER: 30,946
;; REFERENCE/DOCKET NUMBER: LUD 5323.2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 688-9200
;; TELEFAX: (212) 838-3884
;; INFORMATION FOR SEQ ID NO: 16:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 540 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; US-08-602-039-16

Query Match 14.1%; Score 110; DB 3; Length 540;
Best Local Similarity 63.9%; Pred. No. 3.6e-25;
Matches 216; Conservative 0; Mismatches 115; Indels 7; Gaps 3;

QY 447 CCCCCAAAAGAACACAGCAGCTGAAAGTCGGATCCTACACCTGGCGCAGCAGACAGAA 506
DB 203 CCTGAAGAGGGGAACACAGCACTCAAGCTCAGGATCCTGCAGCTGCTCAGGAG---GGA 259
QY 507 GAAGATCAGGATACAGCTGAGATCCCAGTCCGACATGGAAGGTGATCTCCAGAGAGCTG 566
DB 260 GAGGATGAGGAGGATCTGCAAGTCAAGGGCCGAAGCCTGAAGCTGATAGCAGGAACAG 319
QY 567 CATCAGTCAAAACACCGGGGATAATCTGGATTTGGTTCCGCGCTCAAGGTGAAGATAAT 626
DB 320 GGTACCCACAGACTGGGTGTGAGTGTGAAGAT---GGTCTGATGGCAGGAGATGGACCC 378
QY 627 ACCTTAAGAGGAACACTGTAAATGCCAGAGCAGGTGAAGAGCAACCAACCAAGTTTAAT 686
DB 379 GCCAAATCCAGAGGAGGTGAAACGCCCTGAAGAGGTTGAAAGCAATCACAGTGTAAAA 438
QY 687 GAAGCAGAGCTGAACACCAACCAAGCTGGTTTATATAGATATTG---ACTTTAACTAT 743


```
QY 507 GAAGATCAGGATACAGCTGAGATCCAGTCCGCGACATGGAAGGTGATCTGCAAGAGCTG 566
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 252 GAGGATGAGGAGCATCTGCAGGTCAAGGGCCGAAGCTCAAGCTGATAGCCAGGAACAG 311
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 567 CATCAGTCAACACCCGGGATAACTCTGGATTGGGTTCCGGCGTCAAGGTGAAGATAAT 626
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 312 GGTCAACCACAGATGGGTGTGAGTGTGAAGAT--GGTCCTGATGGCGGAGGATGGACCC 370
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 627 ACCTAAAGAGAACACTGTAAATGCCAGAGCAGGTGAAGAGCAACACCAAGTTTAAAT 686
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 371 GCCAAATCCAGAGAGGTGAACGCCCTGAAGAGGTGAAAAGCAATCACAGTCTTAAAA 430
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 687 GAAGACAAGCTGAACACCAAGCTGGTTTATATATAGATATTG---ACTTAAACTAT 743
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 431 GAAGCAGCTTGAATGATCAGGCTGCTCTATGTTGGAATTTGTTCAATTAATTTCT 490
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 744 CTCATAAAGTTTGCAGCTTTCACCAAAAAA 781
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 491 CCCAATAAGCTTTACAGCCTTCTGCAAGAAAAA 528
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
US-08-669-161A-17
; Sequence 17, Application US/08669161A
; Patent No. 6013481
; GENERAL INFORMATION:
; APPLICANT: DeBacker, Olivier; Van den Eynde,
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/669,161A
; FILING DATE: 24-June-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/531,662
; FILING DATE: 21-September-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/370,648
; FILING DATE: 10-January-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/250,162
; FILING DATE: 27-May-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-July-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6013481man D.
; REGISTRATION NUMBER: 30,946
; TELEPHONE: (212) 688-3884
; TELEFAX: (212) 688-9200
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
US-08-669-161A-17
Query Match 13.9%; Score 108.4; DB 3; Length 532;
Best Local Similarity 63.6%; Pred. No. 1.1e-24;
Matches 215; Conservative 0; Mismatches 116; Indels 7; Gaps 3;

QY 447 CCCCCAAAAGAAAGAACAGCAGCTGAAAGTGGGATCTTACACCTGGGCGACAGACAGAA 506
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 195 CCTGAAGAAGGGAAACAGCACTCAACGTCCAGCTCCTGCAGCTGCTCAGGAG---GGA 251
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 507 GAAGATCAGGATACAGCTGAGATCCCAGTCCGCGACATGGAAGGTGATCTGCAAGAGCTG 566
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 252 GAGGATGAGGAGCATCTGCAGGTCAAGGGCCGAAGCTGATAGCCAGGAACAG 311
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 567 CATCAGTCAACACCCGGGATAAATCTGGATTGGGTTCCGGCGTCAAGGTGAAGATAAT 626
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 312 GGTCAACCACAGACTGGGTGTGAGTGTGAAGAT--GGTCCTGATGGCGGAGGATGGACCC 370
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 627 ACCTAAAGAGAACACTGTAAATGCCAGAGCAGGTGAAGAGCAACACCAAGTTTAAAT 686
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 371 GCCAAATCCAGAGAGGTGAAAAGCCCTGAAGAGGTGAAAAGCAATCACAGTCTTAAAA 430
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 687 GAAGACAAGCTGAACACCAAGCTGGTTTATATATAGATATTG---ACTTAAACTAT 743
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 431 GAAGCAGCTTGAATGATCAGGCTGCTCTATGTTGGAATTTGTTCAATTAATTTCT 490
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 744 CTCATAAAGTTTGCAGCTTTCACCAAAAAA 781
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 491 CCCAATAAGCTTTACAGCCTTCTGCAAGAAAAA 528
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
US-08-602-039-17
; Sequence 17, Application US/08602039
; Patent No. 6069001
; GENERAL INFORMATION:
; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
; APPLICANT: Boon-Falleur, Thierry;
; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,039
; FILING DATE: 15-FEB-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/370,648
; FILING DATE: 10-JANUARY-1995
; APPLICATION NUMBER: 08/250,162
; FILING DATE: 27-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-JULY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6069001man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5323.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
```

```
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-602-039-17

Query Match          13.9%; Score 108.4; DB 3; Length 532;
Best Local Similarity 63.6%; Pred. No. 1.1e-24;
Matches 215; Conservative 0; Mismatches 116; Indels 7; Gaps 3;

Qy 447 CCCAAAAGAGAACCCAGCAGCTGAAAGTCGGGATCCTACACCTGGCGACAGCAGAGAA 506
   || || || || || || || || || || || || || || || || || || || || ||
Db 195 CCGTGAAGAGGGGACACCAAGCACTCAACGTCAGGATCTCGCAGCTGCTCAGGAG---GGA 251

Qy 507 GAAGATCAGGATACAGCTGAGATCCAGTCGCGACATGGAAGGTGATCTGCAAGAGCTG 566
   || || || || || || || || || || || || || || || || || || || || ||
Db 252 GAGGATGAGGAGCAGCTGCGAGGTCAAGGGCGGAAGCCTGAAGCTGATAGCCAGGAACAG 311

Qy 567 CATCAGTCAAAACACCGGGGATAAATCTGGATTTGGGTTCCGCGCTCAAGGTGAAGATAAT 626
   || || || || || || || || || || || || || || || || || || || || ||
Db 312 GGTCAACCCACAGACTGGGTGTGAGTGTGAAGAT-GGTCTGTATGGCGAGAGATGGACCC 370

Qy 627 ACCTAAGAGGACACTGTAAATGCCAGACGAGTGAAGAGCAACCAAGTTTAAT 686
   || || || || || || || || || || || || || || || || || || || || ||
Db 371 GCCAATCCAGAGGAGGTGAAACGCCCTGAAGAGGTGAAAGCAATCACAGTGTAAAA 430

Qy 687 GAAGACAAGCTGAAACAACGCAAGCTGGTGTATATATAGATATTG---ACTTAAACTAT 743
   || || || || || || || || || || || || || || || || || || || || ||
Db 431 GAAGCACGTTGAAATGATGACGGCTGCTCTATCTTGAAATTTGTTCAATTAATTTCT 490

Qy 744 CTCATAAAGTTTTCACGCTTTCACCAAAAAA 781
   || || || || || || || || || || || || || || || || || || || || ||
Db 491 CCCAATAAGCTTTACAGCCTTCTGCAAGAGAAAAA 528

RESULT 13
US-08-370-648-14
; Sequence 14, Application US/08370648
; Patent No. 5648226
; GENERAL INFORMATION:
; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
; APPLICANT: Boon-Falleur, Thierry;
; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,648
; FILING DATE: 10-JANUARY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/250,162
; FILING DATE: 27-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-JULY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5648226man D.
; REGISTRATION NUMBER: 30,946
```

```
; REFERENCE/DOCKET NUMBER: LUD 5323.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 538 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-370-648-14

Query Match          13.9%; Score 108.4; DB 1; Length 538;
Best Local Similarity 63.6%; Pred. No. 1.1e-24;
Matches 215; Conservative 0; Mismatches 116; Indels 7; Gaps 3;

Qy 447 CCCAAAAGAGAACCCAGCAGCTGAAAGTCGGGATCTACACCTGGCGACAGCAGACAA 506
   || || || || || || || || || || || || || || || || || || || || ||
Db 201 CCGTGAAGAGGGGACACCAAGCACTCAACGTCAGGATCTCGCAGCTGCTCAGGAG---GGA 257

Qy 507 GAAGATCAGGATACAGCTGAGATCCAGTCGCGACATGGAAGGTGATCTGCAAGAGCTG 566
   || || || || || || || || || || || || || || || || || || || || ||
Db 258 GAGGATGAGGAGCAGCTGCGAGGTCAAGGGCGGAAGCCTGAAGCTCATAGCCAGGAACAG 317

Qy 567 CATCAGTCAAAACACCGGGGATAAATCTGGATTTGGGTTCCGCGCTCAAGGTGAAGATAAT 626
   || || || || || || || || || || || || || || || || || || || || ||
Db 318 GGTCAACCCACAGACTGGGTGTGAGTGTGAAGAT-GGTCTGTATGGCGAGAGATGGACCC 376

Qy 627 ACCTAAGAGGAGCACTGTAAATGCCAGACGAGTGAAGAGCAACCAAGTTTAAT 686
   || || || || || || || || || || || || || || || || || || || || ||
Db 377 GCCAATCCAGAGGAGGTGAAACGCCCTGAAGAGGTGAAAGCAANTCACAGTGTAAAA 436

Qy 687 GAAGACAAGCTGAAACAACGCAAGCTGGTGTATATATAGATATTG---ACTTAAACTAT 743
   || || || || || || || || || || || || || || || || || || || || ||
Db 437 GAAGCACGTTGAAATGATGACGGCTGCTCTATCTTGAAATTTGTTCAATTAATTTCT 496

Qy 744 CTCATAAAGTTTTCACGCTTTCACCAAAAAA 781
   || || || || || || || || || || || || || || || || || || || || ||
Db 497 CCCAATAAGCTTTACAGCCTTCTGCAAGAGAAAAA 534

RESULT 14
US-08-531-662B-14
; Sequence 14, Application US/08531662B
; Patent No. 5858689
; GENERAL INFORMATION:
; APPLICANT: Van der Bruggen, Pierre;
; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
; APPLICANT: Boon-Falleur, Thierry;
; TITLE OF INVENTION: Isolated, Truncated Nucleic
; TITLE OF INVENTION: Acid Molecules Which Code For GAGE Tumor
; TITLE OF INVENTION: Rejection Antigen, The Tumor Rejection Antigen,
; TITLE OF INVENTION: and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531,662B
; FILING DATE: 21-September-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/370,648
; FILING DATE: 10-January-1995
```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/250,162
; FILING DATE: 27-May-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-July-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5858689man D.
; REGISTRATION NUMBER: 30,946
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 538 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-531-662B-14

Query Match 13.9%; Score 108.4; DB 2; Length 538;
Best Local Similarity 63.6%; Pred. No. 1.1e-24;
Matches 215; Conservative 0; Mismatches 116; Indels 7; Gaps 3;

QY 447 CCCCAAAAGAGAACCCAGCAGCTGAAAGTCGGGATCCTACACCTGGCGCAGCAGACAGAA 506
   || || || || || || || || || || || || || || || || || || || || || ||
Db 201 CCTGAAGAGGGGAGCCAGCACTCAACGTCAGGATCCTGCGAGCTGCTCAGGAG--GGA 257

QY 507 GAAGATCAGATACAGCTGAGATCCAGTCGCGACATGGAAGTGATCTGCAAGAGCTG 566
   || || || || || || || || || || || || || || || || || || || || || ||
Db 258 GAGGATGAGGAGGACATCTGCAAGTCGAGGTCGAAGGCGGAGGCTGAAGCTCATAGCCAGGAACAG 317

QY 567 CATCAGTCAAAACACCGGGGATAAATCTGGATTGCGTCCGCGTCAAGGTCAAGATAAT 626
   || || || || || || || || || || || || || || || || || || || || || ||
Db 318 GGTCAACCCACAGACTGGGTGTGAGTGTGAAGAT-GGTCTGATGGCAGGAGATGGACCC 376

QY 627 ACCTAAGAGGACACTGTAAATGCCAGACAGGTGAAGAGCAACACCAAGTTTAAT 686
   || || || || || || || || || || || || || || || || || || || || || ||
Db 377 GCCAAATCCAGAGGAGGTGAAACGCTGAAGAAGGTGAAAGCAATCACAGTGTAAAA 436

QY 687 GAAGACAAGCTGAACAACGCAAGCTGTTTATATATAGATATTG--ACTTAAACTAT 743
   || || || || || || || || || || || || || || || || || || || || || ||
Db 437 GAAGACACGTTGAATGATGCGAGGCTGCTCTATGTTGAAATTTGTTCAATTAATTTCT 496

RESULT 15
US-08-669-161A-14
; Sequence 14, Application US/08669161A
; Patent No. 6013481
; GENERAL INFORMATION:
; APPLICANT: DeBacker, Olivier; Van den Eynde,
; APPLICANT: Benoit; Boon-Falleur, Thierry
; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
; COMPUTER: IBM ps/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/669,161A
; FILING DATE: 24-June-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/531,662
; FILING DATE: 21-September-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/370,648
; FILING DATE: 10-January-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/250,162
; FILING DATE: 27-May-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-July-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6013481man D.
; REGISTRATION NUMBER: 30,946
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 538 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-669-161A-14

Query Match 13.9%; Score 108.4; DB 3; Length 538;
Best Local Similarity 63.6%; Pred. No. 1.1e-24;
Matches 215; Conservative 0; Mismatches 116; Indels 7; Gaps 3;

QY 447 CCCCAAAAGAGAACCCAGCAGCTGAAAGTCGGGATCCTACACCTGGCGCAGCAGACAGAA 506
   || || || || || || || || || || || || || || || || || || || || || ||
Db 201 CCTGAAGAGGGGAGCCAGCACTCAACGTCAGGATCCTGCGAGCTGCTCAGGAG--GGA 257

QY 507 GAAGATCAGGATACAGCTGAGATCCAGTCGCGACATGGAAGTGATCTGCAAGAGCTG 566
   || || || || || || || || || || || || || || || || || || || || || ||
Db 258 GAGGATGAGGAGGACATCTGCAAGTCGAGGTCGAAGGCGGAGGCTGAAGCTCATAGCCAGGAACAG 317

QY 567 CATCAGTCAAAACACCGGGGATAAATCTGGATTGCGTCCGCGTCAAGGTCAAGATAAT 626
   || || || || || || || || || || || || || || || || || || || || || ||
Db 318 GGTCAACCCACAGACTGGGTGTGAGTGTGAAGAT-GGTCTGATGGCAGGAGATGGACCC 376

QY 627 ACCTAAGAGGAGCACTGTAAATGCCAGACAGGTGAAGAGCAACACCAAGTTTAAT 686
   || || || || || || || || || || || || || || || || || || || || || ||
Db 377 GCCAAATCCAGAGGAGGTGAAACGCTGAAGAAGGTGAAAGCAATCACAGTGTAAAA 436

QY 687 GAAGACAAGCTGAACAACGCAAGCTGTTTATATATAGATATTG--ACTTAAACTAT 743
   || || || || || || || || || || || || || || || || || || || || || ||
Db 437 GAAGACACGTTGAATGATGCGAGGCTGCTCTATGTTGAAATTTGTTCAATTAATTTCT 496

RESULT 16
US-08-602-039-14
; Sequence 14, Application US/08602039
; Patent No. 6069001
; GENERAL INFORMATION:
; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch

```


[illegible]

Best Local Similarity 63.6%; Pred. No. 1.2e-24;
Matches 215: Conservative 0: Mismatches 116: Indels 7: Gaps 3:

[illegible]


```
Db 319 GACTCCACCTAATCCTAAGCATGCTAAGCATAAAGAACGAGAGATGGCAGCCATAAGT 378
QY 681 TTAATTAAGACACAGCTGAACAACGCA---AGCTGGTTTTATATTACATATTTGACTTA 737
Db 379 TAAAGAAGACAAAGCTGAAGCTACACACATGGCTGATGTACATTTGAAATGTGACTGA 438
QY 738 AACTATCTCAAT 749
Db 439 AAATTTGAAAT 450

RESULT 22
US-08-669-161A-29
; Sequence 29, Application US/08669161A
; Patent No. 6013481
; GENERAL INFORMATION:
; APPLICANT: DeBacker, Olivier; Van den Eynde,
; APPLICANT: Benoit; Boon-failleu, Thierry
; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/669,161A
; FILING DATE: 24-June-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/531,662
; FILING DATE: 21-September-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/370,648
; FILING DATE: 10-January-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/250,162
; FILING DATE: 27-May-1994
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-July-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No 6013481man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5443
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11461 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-669-161A-29

Query Watch 9.7%; Score 75.4; DB 3; Length 11461;
Best Local Similarity 61.0%; Pred. No. 1.9e-13;
Matches 236; Conservative 0; Mismatches 126; Indels 25; Gaps 6;

QY 28 GGTCCCTGAGGTCTGATTCTTTCCCGCTACTGACACACGCGGGGTAGGTCCACAGGCA 87
Db 1134 GGTTCCTGCGGTCGGGACTCTTTTCCCTACTGAGATTCATCTGTTAGTGTGCAGGCC 1193
```

```
QY 88 GATCCAACCTGGAGTGTGAAGTGTGAGTGAGAGTGAAGAGAACCCAGCAGGCTTCGGGAGG 147
Db 1194 AGTCATCCCGGGGCTGAAGTGTGAGTGAGAGTGAAGAGGCCCTCGGGTGGTTCAGGCGG 1253
QY 148 ----GTTGTGTGGTCTGAGTCTGAGTGTGAGAGTGAAGAGCCCTCGAAGT-----CGTCGTCC 196
Db 1254 GTCCGTTCTCTGTGTGTGGCCCTCCGAGGGAGAAAGGCCACGAGGTACGTACCTCCTTAC 1313
QY 197 CTCTCATCGGTGCCACGCC--CATGGACCTTCTTGTCTCGTCACGGCCATAAATAGGGAG 255
Db 1314 CTTTCACAGGTGCGAGGCCACCGCGGCTTCGTGGTCTGAAGGGGCTGGACGGGGAG 1373
QY 256 GAAGGAGGGCCGAGGAGTGA-----GGGGCTCAGCGGAAGCTGGG----GTGCTGT 303
Db 1374 GAAGGTGGGCGCTGGAGGGGAGGCTGTTCAGGGGCTCAGGTGAAGACGCGGTGAGTGTGT 1433
QY 304 TGGGGGTATCCGAGTCTCTAGAAGCACCTTGAAGCCCGGACAGCAAGATTCTGGACTCCCCAG 363
Db 1434 TGGGGGGATGAAGTCTCC--GAGGTGCGGGGATCCCCCGACACAGGCGAGATCCCTTGA 1492
QY 364 ACGGACACGAGAGGAGGCGGCGCATGAG 390
Db 1493 ATGGGCCCGGGGCGGAGCGGGCG 1519

RESULT 23
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZapt-F1s
US-08-232-463-14
```

Query Match 8.7%; Score 68; DB 1; Length 7218;
Best Local Similarity 3.6%; Pred. NO. 3.2e-11;
Matches 14; Conservative 232; Mismatches 142; Indels 0; Gaps 0;
QY 340 GACAGAGATTCTGGACTCCCGACGAGGACGAGGAGGCGCATGACGACACACA 399
DB 1452 GATAGAAGATTGTTACRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 1393
QY 400 CAAACACACACACACACCCAGTCCAGGACCCAGTAATGAGAGGCCCAAAAGAG 459
DB 1392 RNR 1333
QY 460 AACCCAGCAGTGAAGTCCGGATCTACACCTGGCAGCAGACAGAGAAGATCAGGATA 519
DB 1332 RNR 1273
QY 520 CAGCTGAGATCCAGTCCGCGACATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACA 579
DB 1272 RNR 1213
QY 580 CGGGGATAAACTCGATTGGTTCGCGCTCAAGGTCAAGTCAAGATAATACCTAAAGAGAA 639
DB 1212 RNR 1153
QY 640 CACTGTAAATGCCAGACGAGTGAAGACCAACACAGATTAAATGAAGACAAAGCTGA 699
DB 1152 RNR 1093
QY 700 ACAACGCAAGCTGTTTATATTAGAT 727
DB 1092 RNR 1065

RESULT 24
US-08-370-648-1
Sequence 1, Application US/08370648
Patent No. 5648226
GENERAL INFORMATION:
APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
APPLICANT: Boon-Falleur, Thierry;
TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
TITLE OF INVENTION: Molecules, Which Code For GAGE Tumor Rejection Antigen,
TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/370, 648
FILING DATE: 10-JANUARY-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/250, 162
FILING DATE: 27-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/096, 039
FILING DATE: 22-JULY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5648226man D.
REGISTRATION NUMBER: 30, 946
REFERENCE/DOCKET NUMBER: LUD 5323.2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 646 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-370-648-1
Query Match 7.5%; Score 58.4; DB 1; Length 646;
Best Local Similarity 70.5%; Pred. NO. 8.5e-09;
Matches 93; Conservative 0; Mismatches 36; Indels 3; Gaps 1;
QY 653 CAGAAGCAGGTGACAGACCAACCAAGTTTAAATGAAGACAGCTGAACACGCAAGCT 712
DB 511 CTGTGGCATGTGAAGGGCAATCACAGTGTAAAGAGACATGCTGAAATGTTGAGGCT 570
QY 713 GTTTATATTAGATATTG---ACTTAACTATCTCAATAAAGTTTTCAGCTTTTCACC 769
DB 571 GCTCCTATGTTGGAATAATCTTCATTGAAGTCTCCCAATAAAGCTTTTACAGCCTTCTGC 630
QY 770 AAAAAA 781
DB 631 AAGAAAAA 642
RESULT 25
US-08-531-662B-1
Sequence 1, Application US/08531662B
Patent No. 5858689
GENERAL INFORMATION:
APPLICANT: Van der Bruggen, Pierre;
APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
APPLICANT: Boon-Falleur, Thierry;
TITLE OF INVENTION: Isolated, Truncated Nucleic
TITLE OF INVENTION: Acid Molecules Which Code For GAGE Tumor
TITLE OF INVENTION: Rejection Antigen, The Tumor Rejection Antigen,
TITLE OF INVENTION: and Uses Thereof
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/531,662B
FILING DATE: 21-September-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/370,648
FILING DATE: 10-January-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/250,162
FILING DATE: 27-May-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/096,039
FILING DATE: 22-July-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5858689man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5323.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:


```
QY 690 GACAAGCTGAACAACGCAAGCTGTTTATATTAGATATTTGACTTAAACTATCTCAAT 749
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5162 CAGAAACCTTTACAAATATCTATTACATAAGCAATTTGTCATAGAAATTATACAAGT 5103
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 750 AAGTTTTCAGCTTTCACCAAAAAA 781
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5102 AAATTATACAATATTTAAGCAGCAGCAACATA 5071
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 33
US-08-954-668-18/c
; Sequence 18, Application US/08954668
; Patent No. 6172200
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Eliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/954,668
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36709
; REFERENCE/DOCKET NUMBER: SUV-003.06
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5288 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-954-668-18

Query Match 4.7%; Score 36.8; DB 4; Length 5288;
Best Local Similarity 52.6%; Pred. No. 0.2; Mismatches 0; Gaps 0;
Matches 80; Conservative 0;

QY 630 TAAAGAGGAACACTGTAAATGCCAGAGCAGGTGAAGACCAACCAAGTTTAAATGAA 689
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5222 TTAACAGTAACATTTCTACTACACAGGTTGTGATATGCAAAATTTAAATATTTTAA 5163
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 690 GACAAGCTGAACAACGCAAGCTGTTTATATTAGATATTTGACTTAAACTATCTCAAT 749
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5162 CAGAAACCTTTACAAATATCTATTACATAAGCAATATTTGTCATAGAAATTATACAAGT 5103
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 750 AAGTTTTCAGCTTTCACCAAAAAA 781
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5102 AAATTATACAATATTTAAGCAGCAGCAACATA 5071
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 34
PCT-US95-13233-18/c
; Sequence 18, Application PC/TUS9513233
; GENERAL INFORMATION:
```

```
; APPLICANT: THE BOARD OF TRUSTEES OF TEH LELAND STANFORD JUNIOR UNIVERSITY
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13233
; FILING DATE: 06-OCT-1990
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: a60190-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5288 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
PCT-US95-13233-18

Query Match 4.7%; Score 36.8; DB 5; Length 5288;
Best Local Similarity 52.6%; Pred. No. 0.2; Mismatches 0; Gaps 0;
Matches 80; Conservative 0;

QY 630 TAAAGAGGAACACTGTAAATGCCAGAGCAGGTGAAGACCAACCAAGTTTAAATGAA 689
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5222 TTAACAGTAACATTTCTACTACACAGGTTGTGATATGCAAAATTTAAATATTTTAA 5163
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 690 GACAAGCTGAACAACGCAAGCTGTTTATATTAGATATTTGACTTAAACTATCTCAAT 749
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5162 CAGAAACCTTTACAAATATCTATTACATAAGCAATATTTGTCATAGAAATTATACAAGT 5103
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 750 AAGTTTTCAGCTTTCACCAAAAAA 781
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5102 AAATTATACAATATTTAAGCAGCAGCAACATA 5071
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 35
US-08-920-422-17/c
; Sequence 17, Application US/08920422A
; Patent No. 6255473
; GENERAL INFORMATION:
; APPLICANT: Vitek, Michael P.
; APPLICANT: Mitsuda, No. 6255473iaki
; APPLICANT: Roses, Allen D.
; TITLE OF INVENTION: Presenilin-1 Gene Promoter
; FILE REFERENCE: VITEKPRESENTILIN
; CURRENT APPLICATION NUMBER: US/08/920,422A
; CURRENT FILING DATE: 1997-08-29
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 17
; LENGTH: 48974
; TYPE: DNA
; ORGANISM: Mus musculus
US-08-920-422-17

Query Match 4.6%; Score 36; DB 4; Length 48974;
```

Best Local Similarity 53.6%; Pred. No. 1.4; Mismatches 0; Gaps 0;
Matches 75; Conservative 0; Indels 65; Indels 0; Gaps 0;
QY 372 AGGAGAGGAGCGGATGAGCGACACACACAAACACAGAACACACAGCCAGTCCCGAGGAG 431
Db 44737 AGCCATGACGCACTCACACACACACACACACACACACACTCACACACACACA 44678
QY 432 CCAGTAAATGAGAGCGCCCAAAAGAAAGAACACAGCAGCTCAAAGTCGGGATCCTACACCT 491
Db 44677 CACACACACACACACACACACACACACACAGTGGTGGAGATTAAACTCAGCCTCTTATGTTT 44618
QY 492 GGGCAGCAGACAGAGAAGA 511
Db 44617 GTGTGGCAACACCTGAATA 44598

RESULT 36
US-08-804-227C-1/c
; Sequence 1, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: Dehoff, Bradley S.
; APPLICANT: Kuhnstoss, Stuart A.
; APPLICANT: Rostock, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII(DOS) Text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas, G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43280 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 816..14234
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 14351..19945
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 20010..31199
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31232..36067
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36249..41774
US-08-804-227C-1

Query Match 4.4%; Score 34; DB 2; Length 43280;
Best Local Similarity 50.6%; Pred. No. 5.4; Mismatches 0; Indels 80; Gaps 0;
Matches 82; Conservative 0; Indels 80; Gaps 0;
QY 239 CGCCCAATACTAGGGAGAGAGGCCGCGAGTCCGAGGCGGCTCAGGCGGAAGCTGGGGT 298
Db 18218 CGCCCAATACGCGGAAGGGGGTCCGGTGTCCGCGAGCGCGGAGGCTCGGGGT 18159
QY 299 GCTGTTGGGGGTATCCGAGTCCGAGAGACACCTGGGAACCCCGACAGAGATTCTGGACTC 358
Db 18158 GGTGCGGGGTCTCCCGGGTAAACAGGTCCCGGGCGCGCGCCACCGCCAGGTGTGGT 18099
QY 359 CCAGACGCGGACACGAGGAGGACGCGATCAGCGACACACAC 400
Db 18098 CCGGGCGGTTCGGTTCGCGGAGCGCGTCCAGCGCACCC 18057

RESULT 37
US-07-707-367-1
; Sequence 1, Application US/07707367
; Patent No. 5196316
; GENERAL INFORMATION:
; APPLICANT: Iwasaki, Yasuno
; APPLICANT: Shimoi, Hiroko
; APPLICANT: Suzuki, Kenji
; APPLICANT: Ghisalba, Oreste
; APPLICANT: Nishikawa, Yoshiki
; APPLICANT: Kawahara, Takashi
; APPLICANT: Kangawa, Kenji
; TITLE OF INVENTION: No. 5196316el Enzyme and DNA Coding Therefor
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/707,367
; FILING DATE: 19910530
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 141678/90
; FILING DATE: 01-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 210535/90
; FILING DATE: 10-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 329911/90
; FILING DATE: 30-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Villamizar, JoAnn
; REGISTRATION NUMBER: 30,598
; REFERENCE/DOCKET NUMBER: 4-18110/A/CGJ 44
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914)785-7120
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3383 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (recombinant)
; ORIGINAL SOURCE:
; ORGANISM: Xenopus laevis
; INDIVIDUAL ISOLATE: DNA encoding protein AE-III,

```
; INDIVIDUAL ISOLATE: precursor to PHL enzyme
; IMMEDIATE SOURCE:
; CLONE: PAB-III-202-4
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31..2835
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1177..2145
US-07-707-367-1
```

```
Query Match 4.2%; Score 33; DB 1; Length 3383;
Best Local Similarity 54.5%; Pred. No. 2.5;
Matches 66; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 443 AGAGCCCAAAAGAAAGAACAGCAGCTGAAAGTCGGGATCCTACACCTGGGGCAGCAGAC 502
      |||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2316 AGAAGCAAAATGTTGTGCAAGAGATCAATGCTGGGGTGCTACACAAGAGAGACGAA 2375

QY 503 AGAAGAAGATCAGGATACAGCTGAGATCCCAAGTCGGGACATGGAAGTGATCTGCAAGA 562
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2376 TCTTGCAAGAGAGTAGTGTGGGTGCTCTACACAGAGAGAGAGAGTGTGTGCAAGA 2435

QY 563 G 563
DB 2436 G 2436
```

```
RESULT 38
US-08-487-113D-118
; Sequence 118, Application US/08487113D
; Patent No. 5837822
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Vazeux, Rosemay
; TITLE OF INVENTION: ICAM-Related Materials and Methods
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,113D
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,754
; FILING DATE: 05-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/102,852
; FILING DATE: 05-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/009,266
; FILING DATE: 22-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/894,061
; FILING DATE: 05-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/889,724
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,689
; FILING DATE: 27-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: NO. 5837822and, Greta E.
```

```
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 32744
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 118:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1693 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1587
US-08-487-113D-118

Query Match 4.2%; Score 32.8; DB 2; Length 1693;
Best Local Similarity 56.5%; Pred. No. 1.9;
Matches 61; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 301 TGTGGGGGTATCCGAGTCCAGAACGACCTGGAAACCCCGACAGAAAGATTCTGCACCTCCC 360
      |||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 689 TGTTCGGGCTCGAGGCCCAAGTCCACTTGGCGCTGGGGAACCCAGACGCTGAACCTTA 748

QY 361 CAGACGGGACAGGAGAGGGACGCGCATGAGGACACACACAAACACAG 408
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 749 CAGTCGAGAGCCAGCGGGGACACGATCAGTGCCA/AGCCACAGGCGGTAG 796

RESULT 39
US-08-720-420A-118
; Sequence 118, Application US/08720420A
; Patent No. 5989843
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Vazeux, Rosemay
; TITLE OF INVENTION: ICAM-Related Materials and Methods
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720,420A
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,113
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,754
; FILING DATE: 05-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/102,852
; FILING DATE: 05-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/009,266
; FILING DATE: 22-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/894,061
; FILING DATE: 05-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/889,724
```



```

; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,689
; FILING DATE: 27-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Joseph A., Jr.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 33282
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 118:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1693 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1587
; US-08-720-420A-118

Query Match 4.2%; Score 32.8; DB 2; Length 1693;
Best Local Similarity 56.5%; Pred. No. 1.9;
Matches 61; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 301 TGTTCGGGGTATCCGAGTCGCCAGAAAGACACCTTGGAAACCCGACACAGATTCTTGGACTCCC 360
||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 589 TGTTCGGGCTCCGAGGCCCAAGTCCACTTGGCGCTGGGGACACCGACTGACTCTA 748
||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 361 CACAGCGGGACGAGGAGGGAGCGGCATGAGCGACACACACAAACACAG 408
||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 749 CAGTCCGAGAGCCACGGGACACGATCAGTCCGACAGCGACGACGCGTAG 796
||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 40
US-09-085-199B-6
; Sequence 6, Application US/09085199B
; Patent No. 6235879
; GENERAL INFORMATION:
; APPLICANT: Hayden, Michael R.
; APPLICANT: Hackam, Abigail
; APPLICANT: Huq, A.H.M. Mahbubul
; APPLICANT: Chopra, Vikramjit Singh
; APPLICANT: Kalchman, Michael
; TITLE OF INVENTION: Apoptosis Modulators That Interact with the
; HUNTINGTON'S DISEASE GENE
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oppedahl & Larson
; STREET: PO Box 5270
; CITY: Frisco
; STATE: CO
; COUNTRY: USA
; ZIP: 80443-5270
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Kb storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS DOS 5.0
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/085,199B
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Larson, Marina T.
; REGISTRATION NUMBER: 32038
; REFERENCE/DOCKET NUMBER: UBC_P-013US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (970) 668-2050
; TELEFAX: (970) 668-2052

```

```
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 3251
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: human
; FEATURE:
; OTHER INFORMATION: cDNA for Huntington-interacting protein
US-09-085-199B-6

Query Match          4.1%; Score 32.4; DB 4; Length 3251;
Best Local Similarity 54.1%; Pred. No. 3.7;
Matches 66; Conservative 0; Mismatches 56; Indels 0; Gaps

Qy      488 ACCTGGGCAGCAGACACAAGAATCAGGATACAGCTGAGATCCCATCGTGGCGCACATGGA 547
           | ||||| ||||| || ||||| || ||||| ||||| ||||| ||||| ||||| |||||
Db      508 AGCTGTGGGCGGGGAGCGAGACCTGTGGGCGGCCAGAGCCTGTGTGCCGAGACAGA 567
           | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      548 AGGTGATCTGCAGAGCTGCATCAGTCAGTCAAAACACCGGGGATAAATCTGGATTGGGTTCGG 607
           | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      568 GGCGGCGCTGAGCGCGGAGCAGCAGCCGCGCTCCCAGGACGAGGGCGACTTGCGAGGCCG 627
           | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy      608 GC 609
           ||
Db      628 GC 629
```

Search completed: October 17, 2002, 00:58:28
Job time : 102 seconds

C 144	19	2.4	6728	9	AB019004
C 145	19	2.4	6728	9	AB019004
C 146	19	2.4	6728	9	AB019004

c 220	19	2.4	133790	9	AC010280	AC010280 Homo sapi
c 221	19	2.4	134506	9	HS357116	AL021921 Homo sapi
c 222	19	2.4	135144	9	AC093248	AC093248 Homo sapi
c 223	19	2.4	135671	2	AC107529	AC107529 Rattus no
c 224	19	2.4	136613	2	AC084817	AC084817 Rattus sat
c 225	19	2.4	136943	9	AL359842	AL359842 Human DNA
c 226	19	2.4	137063	9	AC021086	AC021086 Homo sapi
c 227	19	2.4	137229	9	AC010877	AC010877 Homo sapi
c 228	19	2.4	137592	9	AC008625	AC008625 Homo sapi
c 229	19	2.4	138329	2	AC093226	AC093226 Homo sapi
c 230	19	2.4	138414	2	AC008798	AC008798 Homo sapi
c 231	19	2.4	138443	9	HS179EL3	AL035603 Human DNA
c 232	19	2.4	138556	2	AC105564	AC105564 Rattus no
c 233	19	2.4	139566	2	AP004161	AP004161 Rattus sat
c 234	19	2.4	141775	2	AP004664	AP004664 Rattus sat
c 235	19	2.4	142001	2	AC092651	AC092651 Homo sapi
c 236	19	2.4	142629	2	AP002479	AP002479 Homo sapi
c 237	19	2.4	143438	9	AC094067	AC094067 Rattus no
c 238	19	2.4	143463	9	AL359312	AL359312 Human DNA
c 239	19	2.4	143542	2	AC108503	AC108503 Rattus no
c 240	19	2.4	143800	2	AC011875	AC011875 Homo sapi
c 241	19	2.4	144233	2	AC068103	AC068103 Homo sapi
c 242	19	2.4	144984	9	HSJ433F14	AL132766 Human DNA
c 243	19	2.4	146300	9	AL157903	AL157903 Human DNA
c 244	19	2.4	146619	2	OSJN00185	AL662980 Rattus sat
c 245	19	2.4	146950	2	AC095100	AC095100 Rattus no
c 246	19	2.4	147925	2	AC108182	AC108182 Fells cat
c 247	19	2.4	148328	2	AC079850	AC079850 Homo sapi
c 248	19	2.4	149870	9	AC024579	AC024579 Homo sapi
c 249	19	2.4	150401	2	AL139425	AL139425 Homo sapi
c 250	19	2.4	150651	9	AL357499	AL357499 Human DNA
c 251	19	2.4	151365	9	AC010382	AC010382 Homo sapi
c 252	19	2.4	151580	2	OSJN00049	AL606605 Rattus sat
c 253	19	2.4	151851	9	AC013731	AC013731 Homo sapi
c 254	19	2.4	152224	2	AC007933	AC007933 Homo sapi
c 255	19	2.4	152689	9	AL158087	AL158087 Human DNA
c 256	19	2.4	153234	2	AC026947	AC026947 Homo sapi
c 257	19	2.4	153929	2	AC011883	AC011883 Homo sapi
c 258	19	2.4	154384	2	AC015514	AC015514 Homo sapi
c 259	19	2.4	155519	2	AC083992	AC083992 Homo sapi
c 260	19	2.4	155907	2	AC105513	AC105513 Rattus no
c 261	19	2.4	156535	2	AC109344	AC109344 Homo sapi
c 262	19	2.4	157000	2	AC026511	AC026511 Homo sapi
c 263	19	2.4	157067	9	AC010163	AC010163 Homo sapi
c 264	19	2.4	157340	2	AC037493	AC037493 Homo sapi
c 265	19	2.4	157571	2	AC068255	AC068255 Homo sapi
c 266	19	2.4	157921	2	AC027726	AC027726 Homo sapi
c 267	19	2.4	158321	2	AC064810	AC064810 Homo sapi
c 268	19	2.4	158665	2	AC106526	AC106526 Rattus no
c 269	19	2.4	158744	2	AC094158	AC094158 Rattus no
c 270	19	2.4	160274	2	AC073411	AC073411 Homo sapi
c 271	19	2.4	160737	9	HS62F24	AL450003 Homo sapi
c 272	19	2.4	160815	2	AP001895	AP001895 Homo sapi
c 273	19	2.4	160952	9	AP003551	AP003551 Homo sapi
c 274	19	2.4	161207	2	AC092819	AC092819 Homo sapi
c 275	19	2.4	161900	9	AC096729	AC096729 Homo sapi
c 276	19	2.4	162150	9	AC093877	AC093877 Homo sapi
c 277	19	2.4	162828	2	AC106579	AC106579 Rattus no
c 278	19	2.4	163167	2	AC097407	AC097407 Rattus no
c 279	19	2.4	163246	2	AC068401	AC068401 Homo sapi
c 280	19	2.4	163347	2	AP001373	AP001373 Homo sapi
c 281	19	2.4	163372	2	AC087744	AC087744 Homo sapi
c 282	19	2.4	164198	2	AC016801	AC016801 Homo sapi
c 283	19	2.4	164285	2	AC025658	AC025658 Homo sapi
c 284	19	2.4	164820	9	AL627443	AL627443 Human DNA
c 285	19	2.4	165025	2	AC055871	AC055871 Homo sapi
c 286	19	2.4	165533	2	AC094484	AC094484 Rattus no
c 287	19	2.4	165616	9	AC073333	AC073333 Homo sapi
c 288	19	2.4	165655	2	AC027605	AC027605 Homo sapi
c 289	19	2.4	166451	9	AC018994	AC018994 Homo sapi
c 290	19	2.4	166585	9	AC078872	AC078872 Homo sapi
c 291	19	2.4	166894	9	CNS05TCQ	AL355836 Human chr
c 292	19	2.4	166941	2	AC073891	AC073891 Homo sapi
c 293	19	2.4	167348	2	AC013782	AC013782 Homo sapi
c 294	19	2.4	167521	30	AC067761	AC067761 Homo sapi
c 295	19	2.4	167564	2	AF267170	AF267170 Homo sapi
c 296	19	2.4	167919	2	AC090193	AC090193 Homo sapi
c 297	19	2.4	167966	2	AC015912	AC015912 Homo sapi
c 298	19	2.4	168285	9	AC008575	AC008575 Homo sapi
c 299	19	2.4	169078	2	AC094135	AC094135 Rattus no
c 300	19	2.4	169162	2	AC012479	AC012479 Homo sapi
c 301	19	2.4	169554	2	AC021379	AC021379 Homo sapi
c 302	19	2.4	170914	3	AC010843	AC010843 Drosophil
c 303	19	2.4	171225	2	AC024431	AC024431 Homo sapi
c 304	19	2.4	171241	2	AC021558	AC021558 Homo sapi
c 305	19	2.4	171368	9	HS095738	U95738 Human chrom
c 306	19	2.4	171390	9	AC027315	AC027315 Homo sapi
c 307	19	2.4	172024	2	AC078971	AC078971 Homo sapi
c 308	19	2.4	172694	2	CNS01DVS	AL136018 Homo sapi
c 309	19	2.4	172802	9	AC078817	AC078817 Homo sapi
c 310	19	2.4	172830	2	AP001320	AP001320 Homo sapi
c 311	19	2.4	173053	2	AC093523	AC093523 Homo sapi
c 312	19	2.4	173308	9	HS859D4	AL035668 Human DNA
c 313	19	2.4	173840	9	AC092445	AC092445 Homo sapi
c 314	19	2.4	173960	9	AC093885	AC093885 Homo sapi
c 315	19	2.4	173962	2	AC024155	AC024155 Homo sapi
c 316	19	2.4	174428	9	AC005821	AC005821 Homo sapi
c 317	19	2.4	174641	2	AC094401	AC094401 Rattus no
c 318	19	2.4	174994	2	AC025306	AC025306 Homo sapi
c 319	19	2.4	175368	9	CNS01DMF	AL137164 Human chr
c 320	19	2.4	175509	2	AC099333	AC099333 Homo sapi
c 321	19	2.4	175574	2	AC084236	AC084236 Mus muscu
c 322	19	2.4	175602	2	AC097194	AC097194 Rattus no
c 323	19	2.4	177393	2	AC079019	AC079019 Homo sapi
c 324	19	2.4	177829	3	AC091127	AC091127 Drosophil
c 325	19	2.4	179289	2	AC024883	AC024883 Mus muscu
c 326	19	2.4	180317	2	AC025590	AC025590 Homo sapi
c 327	19	2.4	180489	2	AC090383	AC090383 Homo sapi
c 328	19	2.4	180590	9	AC106758	AC106758 Homo sapi
c 329	19	2.4	180919	9	AF088219	AF088219 Homo sapi
c 330	19	2.4	181179	9	AL161725	AL161725 Human DNA
c 331	19	2.4	181311	2	AC106862	AC106862 Homo sapi
c 332	19	2.4	181846	2	AC104651	AC104651 Homo sapi
c 333	19	2.4	181888	2	OSJN00129	AL607004 Rattus sat
c 334	19	2.4	181926	2	AC008461	AC008461 Homo sapi
c 335	19	2.4	181911	9	AC025882	AC025882 Homo sapi
c 336	19	2.4	182175	2	AC013362	AC013362 Homo sapi
c 337	19	2.4	182643	2	AP002794	AP002794 Homo sapi
c 338	19	2.4	182948	2	AL671910	AL671910 Mus muscu
c 339	19	2.4	183205	2	AC013638	AC013638 Homo sapi
c 340	19	2.4	183290	2	AC012569	AC012569 Homo sapi
c 341	19	2.4	183461	9	AC006121	AC006121 Homo sapi
c 342	19	2.4	183719	2	AC068842	AC068842 Homo sapi
c 343	19	2.4	184019	2	AC098901	AC098901 Rattus no
c 344	19	2.4	184244	2	AC073282	AC073282 Homo sapi
c 345	19	2.4	184530	2	AC090707	AC090707 Homo sapi
c 346	19	2.4	184906	2	CNS07ECY	AL445363 Homo sapi
c 347	19	2.4	185573	2	AC010994	AC010994 Drosophil
c 348	19	2.4	185679	2	AL611935	AL611935 Mus muscu
c 349	19	2.4	185967	9	CNS01DVR	AL136001 Human chr
c 350	19	2.4	185981	2	AC095185	AC095185 Rattus no
c 351	19	2.4	186464	9	AC092839	AC092839 Homo sapi
c 352	19	2.4	186709	2	AC091192	AC091192 Homo sapi
c 353	19	2.4	186962	9	AP000848	AP000848 Homo sapi
c 354	19	2.4	187617	2	AC084082	AC084082 Homo sapi
c 355	19	2.4	187691	10	AC083816	AC083816 Mus muscu
c 356	19	2.4	188107	9	CNS01DWE	AL137129 Human chr
c 357	19	2.4	189371	2	AC023125	AC023125 Homo sapi
c 358	19	2.4	189897	2	AC104829	AC104829 Homo sapi
c 359	19	2.4	190672	2	AC105890	AC105890 Rattus no
c 360	19	2.4	190843	8	AP002863	AP002863 Rattus sat
c 361	19	2.4	190875	2	AC022255	AC022255 Homo sapi
c 362	19	2.4	191360	2	AL663079	AL663079 Mus muscu
c 363	19	2.4	191770	2	AC093616	AC093616 Homo sapi
c 364	19	2.4	191835	2	AC007595	AC007595 Homo sapi
c 365	19	2.4	192083	2	AC015901	AC015901 Homo sapi

us-09-658-824-808.oligo.rge

[illegible]

c 512	18	2.3	5031	10	AF199332	AF199332 Rattus no	c 585	18	2.3	44583	9	AC000083	AC000083 Homo sapi
c 513	18	2.3	5079	10	AF199325	AF199325 Rattus no	c 586	18	2.3	45255	2	AC101034	AC101034 Mus muscu
c 514	18	2.3	5133	9	HSFLA1A	Y00796 Human mRNA	c 587	18	2.3	45291	2	AC093987	AC093987 Rattus no
c 515	18	2.3	5172	10	AF199324	AF199324 Rattus no	c 588	18	2.3	47507	2	AC017676	AC017676 Drosophill
c 516	18	2.3	5223	6	AX252154	AX252154 Sequence	c 589	18	2.3	47724	8	AC007119	AC007119 Arabidops
c 517	18	2.3	5233	6	AX344520	AX344520 Sequence	c 590	18	2.3	50785	8	AC007169	AC007169 Arabidops
c 518	18	2.3	5233	6	AX348927	AX348927 Sequence	c 591	18	2.3	54526	2	AC087404	AC087404 Homo sapi
c 519	18	2.3	5253	6	AX348931	AX348931 Rattus no	c 592	18	2.3	58423	2	AC087660	AC087660 Homo sapi
c 520	18	2.3	5550	10	AF199331	AF199331 Rattus no	c 593	18	2.3	59026	2	AC101623	AC101623 Mus muscu
c 521	18	2.3	5592	10	AF199330	AF199330 Rattus no	c 594	18	2.3	60318	2	AC110240	AC110240 Mus muscu
c 522	18	2.3	5640	10	AF199332	AF199332 Rattus no	c 595	18	2.3	60610	9	AL591470	AL591470 Human DNA
c 523	18	2.3	5680	6	AX346731	AX346731 Sequence	c 596	18	2.3	61169	9	AP001433	AP001433 Homo sapi
c 524	18	2.3	5788	8	ALPAAAT2A	L25335 Medicago sa	c 597	18	2.3	61623	2	AC108413	AC108413 Mus muscu
c 525	18	2.3	6012	6	AX251974	AX251974 Sequence	c 598	18	2.3	61823	2	AC108413	AC108413 Human DNA
c 526	18	2.3	6012	6	AX344366	AX344366 Sequence	c 599	18	2.3	62493	9	HS193G15	HS193G15 Human DNA
c 527	18	2.3	6012	6	AX348759	AX348759 Sequence	c 600	18	2.3	62493	9	HS193G15	HS193G15 Human DNA
c 528	18	2.3	6113	9	F325326S06	AF325331 Homo sapi	c 601	18	2.3	62612	2	AC107699	AC107699 Mus muscu
c 529	18	2.3	6113	9	F325326S06	AF325331 Homo sapi	c 602	18	2.3	62856	2	AC108314	AC108314 Rattus no
c 530	18	2.3	6962	8	FN062895	U62895 Emericella	c 603	18	2.3	63212	2	AC107153	AC107153 Drosophill
c 531	18	2.3	7226	14	S74099	AB000125 Aspergill	c 604	18	2.3	63767	8	ATAC023912	ATAC023912 Arabidops
c 532	18	2.3	7692	10	AB030013	S74099 Avian myelo	c 605	18	2.3	64582	8	AC027037	AC027037 Oryza sat
c 533	18	2.3	8064	6	AX346288	AB030013 Mus muscu	c 606	18	2.3	64582	8	AC027037	AC027037 Oryza sat
c 534	18	2.3	9646	6	AX346591	AX346288 Sequence	c 607	18	2.3	65493	2	AC099807	AC099807 Homo sapi
c 535	18	2.3	10820	6	AX305711	AX346591 Sequence	c 608	18	2.3	66008	9	AC023190	AC023190 Homo sapi
c 536	18	2.3	10820	6	AX305711	AX305711 Sequence	c 609	18	2.3	66375	2	AC100703	AC100703 Mus muscu
c 537	18	2.3	10951	1	AE007263	U62021 Mus musculu	c 610	18	2.3	66714	2	AC110187	AC110187 Mus muscu
c 538	18	2.3	11015	6	AX280003	AE007263 Sinorhizo	c 611	18	2.3	66966	2	AC026842	AC026842 Homo sapi
c 539	18	2.3	11015	6	AX281190	AX280003 Sequence	c 612	18	2.3	67780	2	AC087275	AC087275 Homo sapi
c 540	18	2.3	11015	6	AX336459	AX281190 Sequence	c 613	18	2.3	68003	2	AC090851	AC090851 Homo sapi
c 541	18	2.3	11262	1	AE004170	AE004170 Vibrio ch	c 614	18	2.3	68723	2	AC019515	AC019515 Drosophill
c 542	18	2.3	11557	9	AF187881	AF187881 Homo sapi	c 615	18	2.3	68991	2	AC102157	AC102157 Mus muscu
c 543	18	2.3	11726	6	AX346964	AX346964 Sequence	c 616	18	2.3	69223	2	AC102441	AC102441 Mus muscu
c 544	18	2.3	12054	6	AX346080	AX346080 Sequence	c 617	18	2.3	69304	2	AC102598	AC102598 Mus muscu
c 545	18	2.3	12307	2	AC024562_4	Continuation (5 of	c 618	18	2.3	69471	2	AC099172	AC099172 Rattus no
c 546	18	2.3	12537	6	AR069468	AR069468 Sequence	c 619	18	2.3	69682	9	AC080049	AC080049 Homo sapi
c 547	18	2.3	12537	6	AR162785	AR162785 Sequence	c 620	18	2.3	69682	9	AC080049	AC080049 Homo sapi
c 548	18	2.3	12599	3	AC024743	AC024743 Caenorhab	c 621	18	2.3	71327	8	AL596452	AL596452 Human DNA
c 549	18	2.3	13131	6	AX356424	AX356424 Sequence	c 622	18	2.3	72596	2	AC010300	AC010300 Arabidops
c 550	18	2.3	13131	6	AX356424	AX356424 Sequence	c 623	18	2.3	74968	6	AB019230	AB019230 Arabidops
c 551	18	2.3	14543	1	AE001691	AE001691 Thermotog	c 624	18	2.3	75611	2	AC102539	AC102539 Mus muscu
c 552	18	2.3	15218	2	AC014602	AC014602 Drosophill	c 625	18	2.3	75611	2	AC102539	AC102539 Mus muscu
c 553	18	2.3	15755	10	MMU20949	U20949 Mus musculu	c 626	18	2.3	77331	9	AL354983	AL354983 Human DNA
c 554	18	2.3	16157	2	AC012931	AC012931 Drosophill	c 627	18	2.3	77331	9	AL354983	AL354983 Human DNA
c 555	18	2.3	17343	2	AC104491	AC104491 Trypanoso	c 628	18	2.3	77380	8	AB010077	AB010077 Arabidops
c 556	18	2.3	17421	6	AX277891	AX277891 Sequence	c 629	18	2.3	78184	9	AC067722	AC067722 Homo sapi
c 557	18	2.3	17421	6	AX323568	AX323568 Sequence	c 630	18	2.3	78485	2	AC027593	AC027593 Homo sapi
c 558	18	2.3	17720	2	AC104495	AC104495 Trypanoso	c 631	18	2.3	79414	2	AC023502	AC023502 Homo sapi
c 559	18	2.3	17986	3	CEL2K75	U23451 Caenorhabdi	c 632	18	2.3	79509	9	AL353769	AL353769 Human DNA
c 560	18	2.3	20029	9	AL589646	AL589646 Human DNA	c 633	18	2.3	79531	9	AL390037	AL390037 Human DNA
c 561	18	2.3	20587	9	HS185E8A	U23451 Caenorhabdi	c 634	18	2.3	80376	2	AC004442	AC004442 Drosophill
c 562	18	2.3	20587	9	HS185E8A	U23451 Caenorhabdi	c 635	18	2.3	80376	2	AC095072	AC095072 Rattus no
c 563	18	2.3	20872	9	AP000222	Z68274 Human DNA S	c 636	18	2.3	81020	8	AB026645	AB026645 Arabidops
c 564	18	2.3	22586	3	AC094018	AP000222 Homo sapi	c 637	18	2.3	81020	8	AB026645	AB026645 Arabidops
c 565	18	2.3	22843	2	AC015232	AC094018 Trypanoso	c 638	18	2.3	81347	8	AB015478	AB015478 Arabidops
c 566	18	2.3	25659	10	AB022160	AC015232 Drosophill	c 639	18	2.3	82912	3	AC004574	AC004574 Human DNA
c 567	18	2.3	29498	8	AB023029	AB022160 Mus muscu	c 640	18	2.3	82912	3	AC004574	AC004574 Human DNA
c 568	18	2.3	29663	2	AC014990	AB023029 Arabidops	c 641	18	2.3	83253	8	AP000383	AP000383 Arabidops
c 569	18	2.3	30616	2	AC018027	AC014990 Drosophill	c 642	18	2.3	83253	8	AP000383	AP000383 Arabidops
c 570	18	2.3	31299	8	SPCC18B5	AC018027 Drosophill	c 643	18	2.3	83989	9	AL135928	AL135928 Human DNA
c 571	18	2.3	32000	2	AC092777	AC018255 Drosophill	c 644	18	2.3	85992	8	ATF2G14	ATF2G14 Arabidops
c 572	18	2.3	33772	2	AF045638	AL109736 S. pombe c	c 645	18	2.3	85992	8	ATF2G14	ATF2G14 Arabidops
c 573	18	2.3	34561	2	AC015420	AF029777 Leishmani	c 646	18	2.3	86554	8	AB005245	AB005245 Arabidops
c 574	18	2.3	35469	8	HS786D3	AF045638 Caenorhab	c 647	18	2.3	86630	9	AB005245	AB005245 Arabidops
c 575	18	2.3	35930	8	AC079288	AC015420 Drosophill	c 648	18	2.3	86630	9	AB005245	AB005245 Arabidops
c 576	18	2.3	38614	2	AC107503	AL023801 Human DNA	c 649	18	2.3	87224	2	AC097775	AC097775 Rattus no
c 577	18	2.3	38865	2	AC101934	AL023801 Human DNA	c 650	18	2.3	87224	2	AC097775	AC097775 Rattus no
c 578	18	2.3	39006	3	AF077531	AC107503 Rattus no	c 651	18	2.3	87420	2	CNS050TL	CNS050TL Human DNA
c 579	18	2.3	40018	9	HS10111	AC101934 Drosophill	c 652	18	2.3	88484	9	AL355794	AL355794 Human DNA
c 580	18	2.3	40331	9	AL590223	AF077531 Caenorhab	c 653	18	2.3	88484	9	AL355794	AL355794 Human DNA
c 581	18	2.3	41015	9	HS590K14	AL109616 Homo sapi	c 654	18	2.3	88484	9	AL355794	AL355794 Human DNA
c 582	18	2.3	41052	9	AC002499	AL590223 Human DNA	c 655	18	2.3	89669	9	HS0340H1	HS0340H1 Human DNA
c 583	18	2.3	41609	3	CEL03A7	AL049198 Human DNA	c 656	18	2.3	91340	9	AC008898	AC008898 Homo sapi
c 584	18	2.3	44571	2	AC104968	AF016451 Caenorhab	c 657	18	2.3	92118	9	AC074384	AC074384 Homo sapi

Thu Oct 17 09:19:15 2002

c 658	18	2.3	92458	9	AL353715	Human DNA	AL353715	Human DNA	AC025170	9	AC025170	Homo sapi
c 659	18	2.3	92705	9	AC006327	Homo sapi	AC006327	Homo sapi	AL109914	9	AL109914	Human DNA
c 660	18	2.3	92741	8	AP004473	Lotus jap	AP004473	Lotus jap	AC002464	9	AC002464	Human BAC
c 661	18	2.3	92918	2	AC020050	Drosophila	AC020050	Drosophila	AL160169	9	AL160169	Human DNA
c 662	18	2.3	93244	2	AC090062	Homo sapi	AC090062	Homo sapi	AC087346	9	AC087346	Homo sapi
c 663	18	2.3	93708	9	AL135934	Human DNA	AL135934	Human DNA	AC010475	9	AC010475	Homo sapi
c 664	18	2.3	93764	2	AC074072	Homo sapi	AC074072	Homo sapi	AL031726	9	AL031726	Human DNA
c 665	18	2.3	93795	2	AC022428	Homo sapi	AC022428	Homo sapi	AC010589	9	AC010589	Homo sapi
c 666	18	2.3	93891	9	AC024192	Homo sapi	AC024192	Homo sapi	AC093510	9	AC093510	Homo sapi
c 667	18	2.3	95481	2	AC093284	Homo sapi	AC093284	Homo sapi	AC008879	9	AC008879	Homo sapi
c 668	18	2.3	95554	2	AC004783	Homo sapi	AC004783	Homo sapi	AL096840	9	AL096840	Human DNA
c 669	18	2.3	96654	9	AC004783	Homo sapi	AC004783	Homo sapi	AC095554	9	AC095554	Rattus no
c 670	18	2.3	96899	8	AC013288	Arabidops	AC013288	Arabidops	AC004976	9	AC004976	Homo sapi
c 671	18	2.3	98002	9	AL353147	Human DNA	AL353147	Human DNA	AC003958	9	AC003958	Homo sapi
c 672	18	2.3	98195	9	AC098045	Homo sapi	AC098045	Homo sapi	AC004457	9	AC004457	Homo sapi
c 673	18	2.3	98595	9	AP001329	Homo sapi	AP001329	Homo sapi	AC004241	9	AC004241	Homo sapi
c 674	18	2.3	98939	2	AC107563	Rattus no	AC107563	Rattus no	AL078583	9	AL078583	Human DNA
c 675	18	2.3	99589	2	AC093299	Homo sapi	AC093299	Homo sapi	AC079038	9	AC079038	Oryza sat
c 676	18	2.3	99637	2	AC096802	Rattus no	AC096802	Rattus no	AC092831	9	AC092831	Homo sapi
c 677	18	2.3	99923	8	F508	Arabidops	F508	Arabidops	AC056664	9	AC056664	Homo sapi
c 678	18	2.3	100000	9	AB020864	Homo sapi	AB020864	Homo sapi	AC097737	9	AC097737	Rattus no
c 679	18	2.3	100000	9	AP000085	Homo sapi	AP000085	Homo sapi	AC068667	9	AC068667	Arabidops
c 680	18	2.3	100000	9	AP000137	Homo sapi	AP000137	Homo sapi	AF165145	9	AF165145	Homo sapi
c 681	18	2.3	100000	9	AP000158	Homo sapi	AP000158	Homo sapi	AC069506	9	AC069506	Homo sapi
c 682	18	2.3	100328	8	AT18115	Arabidops	AT18115	Arabidops	AC005171	9	AC005171	Arabidops
c 683	18	2.3	100786	9	AC013972	Drosophila	AC013972	Drosophila	AC093850	9	AC093850	Homo sapi
c 684	18	2.3	101027	9	AC023160	Homo sapi	AC023160	Homo sapi	AC105528	9	AC105528	Rattus no
c 685	18	2.3	101121	2	AC097166	Homo sapi	AC097166	Homo sapi	AC093178	9	AC093178	Oryza sat
c 686	18	2.3	101453	2	AC098252	Human DNA	AC098252	Human DNA	AC021756	9	AC021756	Mus muscu
c 687	18	2.3	10158	9	AL590005	Human DNA	AL590005	Human DNA	AL356293	9	AL356293	Homo sapi
c 688	18	2.3	101670	2	AC105714	Rattus no	AC105714	Rattus no	AC008143	9	AC008143	Drosophila
c 689	18	2.3	101966	8	AC012563	Arabidops	AC012563	Arabidops	AL445259	9	AL445259	Human DNA
c 690	18	2.3	102183	9	AC027134	Human DNA	AC027134	Human DNA	AL160398	9	AL160398	Human DNA
c 691	18	2.3	102892	9	AL359813	Human DNA	AL359813	Human DNA	AL512649	9	AL512649	Human DNA
c 692	18	2.3	102994	9	AC099083	Rattus no	AC099083	Rattus no	AL121914	9	AL121914	Human DNA
c 693	18	2.3	103157	8	AC011810	Arabidops	AC011810	Arabidops	AC092781	9	AC092781	Oryza sat
c 694	18	2.3	103567	8	YUP812	Arabidops	YUP812	Arabidops	AC099677	9	AC099677	Homo sapi
c 695	18	2.3	103576	8	HS124C6	Homo sapi	HS124C6	Homo sapi	AL606487	9	AL606487	Mouse DNA
c 696	18	2.3	104436	6	HS611N7	Rattus no	HS611N7	Rattus no	AC018487	9	AC018487	Drosophila
c 697	18	2.3	105956	2	AC098919	Homo sapi	AC098919	Homo sapi	AP004928	9	AP004928	Homo sapi
c 698	18	2.3	107139	2	AL360223	Homo sapi	AL360223	Homo sapi	AP003894	9	AP003894	Oryza sat
c 699	18	2.3	109389	9	AC002530	Human DNA	AC002530	Human DNA	AL161732	9	AL161732	Human DNA
c 700	18	2.3	109612	9	AL512503	Human DNA	AL512503	Human DNA	AL136368	9	AL136368	Human DNA
c 701	18	2.3	109813	9	HSJ180E22	Human DNA	HSJ180E22	Human DNA	AC093265	9	AC093265	Homo sapi
c 702	18	2.3	110000	2	AC073604.2	Continuation (3 of	AC073604.2	Continuation (3 of	AP004762	9	AP004762	Oryza sat
c 703	18	2.3	110000	2	AC092202.1	Continuation (2 of	AC092202.1	Continuation (2 of	AJ249895	9	AJ249895	Mus muscu
c 704	18	2.3	110000	2	AC026388.1	Continuation (2 of	AC026388.1	Continuation (2 of	AP004563	9	AP004563	Oryza sat
c 705	18	2.3	110338	3	AC006066	Drosophila	AC006066	Drosophila	AL135916	9	AL135916	Homo sapi
c 706	18	2.3	111122	2	AC094837	Rattus no	AC094837	Rattus no	AP004754	9	AP004754	Oryza sat
c 707	18	2.3	111123	2	AC066604	Homo sapi	AC066604	Homo sapi	AC103044	9	AC103044	Rattus no
c 708	18	2.3	112298	2	AC066370	Homo sapi	AC066370	Homo sapi	AC024510	9	AC024510	Homo sapi
c 709	18	2.3	112390	9	AC016607	Homo sapi	AC016607	Homo sapi	AC004010	9	AC004010	Homo sapi
c 710	18	2.3	112535	2	AC103152	Rattus no	AC103152	Rattus no	AC023197	9	AC023197	Mus muscu
c 711	18	2.3	113367	9	AC008178	Homo sapi	AC008178	Homo sapi	AL138761	9	AL138761	Human DNA
c 712	18	2.3	113501	2	AC011398	Homo sapi	AC011398	Homo sapi	AC090553	9	AC090553	Sus scrof
c 713	18	2.3	114008	2	AC093780	Homo sapi	AC093780	Homo sapi	AC108020	9	AC108020	Homo sapi
c 714	18	2.3	114793	9	AF217796	Human DNA	AF217796	Human DNA	AC104122	9	AC104122	Homo sapi
c 715	18	2.3	115278	9	AL139819	Human DNA	AL139819	Human DNA	AC023467	9	AC023467	Homo sapi
c 716	18	2.3	115468	2	AC011820	Homo sapi	AC011820	Homo sapi	AC011972	9	AC011972	Homo sapi
c 717	18	2.3	115954	9	AC005532	Homo sapi	AC005532	Homo sapi	AC097715	9	AC097715	Homo sapi
c 718	18	2.3	116106	2	AL162611	Homo sapi	AL162611	Homo sapi	AL365366	9	AL365366	Human DNA
c 719	18	2.3	116181	9	AC010681	Homo sapi	AC010681	Homo sapi	AC027566	9	AC027566	Homo sapi
c 720	18	2.3	117840	9	AC010245	Homo sapi	AC010245	Homo sapi	AL353139	9	AL353139	Homo sapi
c 721	18	2.3	118524	9	HSJ442L6	Human DNA	HSJ442L6	Human DNA	AL160276	9	AL160276	Human DNA
c 722	18	2.3	118873	9	AL157878	Human DNA	AL157878	Human DNA	AL139001	9	AL139001	Human DNA
c 723	18	2.3	118955	9	AC068274	Homo sapi	AC068274	Homo sapi	AC016294	9	AC016294	Homo sapi
c 724	18	2.3	120192	2	AC093220	Homo sapi	AC093220	Homo sapi	AC018833	9	AC018833	Homo sapi
c 725	18	2.3	120219	9	AL355540	Human DNA	AL355540	Human DNA				
c 726	18	2.3	120955	9	HUAC002310	Human Chr	HUAC002310	Human Chr				
c 727	18	2.3	121598	9	AC060612	Homo sapi	AC060612	Homo sapi				
c 728	18	2.3	122850	9	AL359744	Human DNA	AL359744	Human DNA				
c 729	18	2.3	122952	2	AL513483	Homo sapi	AL513483	Homo sapi				
c 730	18	2.3	123183	8	AC006841	Arabidops	AC006841	Arabidops				

c 804	18	2.3 146240	2	AC103012	AC103012 Rattus no	877	18	2.3 158892	2	CNS01RI4	AL162871 Homo sapi
805	18	2.3 147114	9	AC008600	AC008600 Homo sapi	c 878	18	2.3 158918	2	AL391806	AL391806 Homo sapi
806	18	2.3 147140	9	AC010226	AC010226 Homo sapi	879	18	2.3 159315	2	AC026918	AC026918 Homo sapi
807	18	2.3 147431	2	AC020996	AC020996 Homo sapi	880	18	2.3 159362	9	AL589741	AL589741 Human DNA
c 808	18	2.3 147700	2	AP002362	AP002362 Homo sapi	881	18	2.3 159409	2	AC092900	AC092900 Homo sapi
809	18	2.3 148041	2	AC021320	AC021320 Homo sapi	c 882	18	2.3 159435	2	AL108696	AL108696 Homo sapi
c 810	18	2.3 148087	2	AC094950	AC094950 Rattus no	c 883	18	2.3 159555	2	AL591710	AL591710 Homo sapi
c 811	18	2.3 148846	2	AC008714	AC008714 Homo sapi	884	18	2.3 159620	2	AP002442	AP002442 Homo sapi
812	18	2.3 148876	9	AC011455	AC011455 Homo sapi	c 885	18	2.3 159838	9	AC073991	AC073991 Homo sapi
813	18	2.3 149114	9	AP001172	AP001172 Homo sapi	c 886	18	2.3 160149	2	AC016380	AC016380 Homo sapi
814	18	2.3 149597	2	AC034271	AC034271 Homo sapi	887	18	2.3 160154	2	AL391086	AL391086 Homo sapi
c 815	18	2.3 149925	2	AC094577	AC094577 Rattus no	c 888	18	2.3 160551	3	AC008320	AC008320 Drosophil
c 816	18	2.3 150263	9	AC073218	AC073218 Homo sapi	889	18	2.3 160877	2	AC026560	AC026560 Homo sapi
817	18	2.3 150397	2	CNS05TD2	AL355923 Homo sapi	c 890	18	2.3 161021	2	AL451068	AL451068 Homo sapi
c 818	18	2.3 150407	2	AC093228	AC093228 Homo sapi	891	18	2.3 161374	2	AC037430	AC037430 Homo sapi
819	18	2.3 150866	2	AC105311	AC105311 Homo sapi	892	18	2.3 161390	9	AC013414	AC013414 Homo sapi
820	18	2.3 150974	2	AL672064	AL672064 Mus muscu	c 893	18	2.3 161507	9	AC046195	AC046195 Homo sapi
821	18	2.3 150989	9	AC091922	AC091922 Homo sapi	c 894	18	2.3 161613	2	AC023947	AC023947 Homo sapi
c 822	18	2.3 151040	2	AC020827	AC020827 Mus muscu	c 895	18	2.3 161885	2	AC026518	AC026518 Homo sapi
823	18	2.3 151164	2	AC109808	AC109808 Homo sapi	896	18	2.3 161990	9	AC024723	AC024723 Homo sapi
c 824	18	2.3 151821	2	AC098213	AC098213 Rattus no	c 897	18	2.3 162269	9	AC000097	AC000097 Homo sapi
825	18	2.3 151894	2	AL389926	AL389926 Homo sapi	c 898	18	2.3 162549	9	AC090802	AC090802 Homo sapi
826	18	2.3 152103	2	AC018992	AC018992 Homo sapi	c 899	18	2.3 162692	2	AC016171	AC016171 Homo sapi
827	18	2.3 152366	9	AL138894	AL138894 Human DNA	c 900	18	2.3 163179	9	AC013549	AC013549 Homo sapi
c 828	18	2.3 152443	3	AC009208	AC009208 Drosophil	c 901	18	2.3 163301	9	CNS01DWB	AL136537 Human chr
c 829	18	2.3 152448	2	AC021391	AC021391 Homo sapi	902	18	2.3 163378	2	AC096002	AC096002 Rattus no
830	18	2.3 152537	9	AC055731	AC055731 Homo sapi	c 903	18	2.3 163538	9	AC012598	AC012598 Homo sapi
c 831	18	2.3 152685	9	AC008818	AC008818 Homo sapi	c 904	18	2.3 163577	9	AL590428	AL590428 Human DNA
832	18	2.3 152782	9	AL132555	AL132555 Human DNA	c 905	18	2.3 163619	9	AL162502	AL162502 Human DNA
c 833	18	2.3 152853	3	AC009389	AC009389 Drosophil	c 906	18	2.3 163988	9	AC007673	AC007673 Homo sapi
834	18	2.3 153048	3	AC007891	AC007891 Drosophil	907	18	2.3 164073	2	AC008032	AC008032 Homo sapi
835	18	2.3 153092	9	AC069475	AC069475 Homo sapi	c 908	18	2.3 164268	9	AC022402	AC022402 Homo sapi
c 836	18	2.3 153337	2	AC095309	AC095309 Rattus no	c 909	18	2.3 164490	2	AC079077	AC079077 Homo sapi
c 837	18	2.3 153607	2	AL355523	AL355523 Homo sapi	c 910	18	2.3 164503	2	AC020630	AC020630 Homo sapi
c 838	18	2.3 153735	9	AC004029	AC004029 Human BAC	c 911	18	2.3 164736	2	AL390855	AL390855 Homo sapi
c 839	18	2.3 153783	9	AL157937	AL157937 Human DNA	912	18	2.3 164741	2	AC027278	AC027278 Mus muscu
840	18	2.3 153830	9	AL513191	AL513191 Human DNA	c 913	18	2.3 164755	2	AC096916	AC096916 Homo sapi
c 841	18	2.3 154018	2	AC062019	AC062019 Homo sapi	c 914	18	2.3 164794	2	AC087361	AC087361 Homo sapi
842	18	2.3 154158	2	AL365257	AL365257 Homo sapi	c 915	18	2.3 164839	8	AP002844	AP002844 Oryza sat
c 843	18	2.3 154242	2	AL442071	AL442071 Homo sapi	c 916	18	2.3 164863	2	AC026517	AC026517 Homo sapi
844	18	2.3 154618	2	AC073403	AC073403 Homo sapi	917	18	2.3 164991	9	AC011088	AC011088 Homo sapi
c 845	18	2.3 154706	2	AC008470	AC008470 Homo sapi	918	18	2.3 165016	9	AC024559	AC024559 Homo sapi
c 846	18	2.3 154887	9	AL137060	AL137060 Human DNA	919	18	2.3 165063	2	AL607025	AL607025 Mus muscu
c 847	18	2.3 155359	9	AP003689	AP003689 Homo sapi	c 920	18	2.3 165082	2	AC106856	AC106856 Homo sapi
c 848	18	2.3 155376	2	AC027026	AC027026 Homo sapi	921	18	2.3 165118	3	AC108488	AC108488 Drosophil
c 849	18	2.3 155420	8	AC025906	AC025906 Oryza sat	c 922	18	2.3 165126	9	CNS01IRHV	AL162511 Human chr
c 850	18	2.3 155685	3	AC098575	AC098575 Drosophil	c 923	18	2.3 165554	2	AC022528	AC022528 Homo sapi
c 851	18	2.3 155688	2	AL662861	AL662861 Danio rer	c 924	18	2.3 165676	2	AC010268	AC010268 Homo sapi
c 852	18	2.3 155742	9	AC006354	AC006354 Homo sapi	925	18	2.3 165730	2	AC018652	AC018652 Homo sapi
c 853	18	2.3 155818	30	AC009768	AC009768 Homo sapi	c 926	18	2.3 165817	2	AC007935	AC007935 Homo s.p.i
c 854	18	2.3 155992	2	AC105602	AC105602 Rattus no	927	18	2.3 165845	3	AC010022	AC010022 Drosop. il
c 855	18	2.3 156074	2	AC020640	AC020640 Homo sapi	c 928	18	2.3 165907	2	AC093802	AC093802 Homo sapi
856	18	2.3 156108	2	AL354944	AL354944 Homo sapi	c 929	18	2.3 166041	9	AC006987	AC006987 Homo sapi
857	18	2.3 156265	9	AC079140	AC079140 Homo sapi	930	18	2.3 166045	2	AC018543	AC018543 Homo sapi
c 858	18	2.3 156284	2	AC094381	AC094381 Rattus no	c 931	18	2.3 166075	2	AC092836	AC092836 Homo sapi
859	18	2.3 156288	2	AC021069	AC021069 Homo sapi	c 932	18	2.3 166206	2	AC016927	AC016927 Homo sapi
860	18	2.3 156600	2	AC109587	AC109587 Homo sapi	c 933	18	2.3 166239	2	AC078965	AC078965 Homo sapi
c 861	18	2.3 156760	2	AC022388	AC022388 Homo sapi	c 934	18	2.3 166253	9	AC008900	AC008900 Homo sapi
862	18	2.3 156970	2	AC018386	AC018386 Homo sapi	935	18	2.3 166266	2	AP001146	AP001146 Homo sapi
863	18	2.3 157017	9	AL162389	AL162389 Human DNA	c 936	18	2.3 166288	9	AL139410	AL139410 Human DNA
864	18	2.3 157051	9	AC011095	AC011095 Homo sapi	c 937	18	2.3 166427	2	AC069593	AC069593 Homo sapi
865	18	2.3 157081	9	AL139186	AL139186 Human DNA	938	18	2.3 166608	7	AC024727	AC024727 Homo sapi
866	18	2.3 157178	9	AC023095	AC023095 Homo sapi	c 939	18	2.3 166637	9	AC008657	AC008657 Homo sapi
867	18	2.3 157296	2	AL359705	AL359705 Homo sapi	940	18	2.3 166642	3	AC008364	AC008364 Drosophil
868	18	2.3 157784	9	AL360232	AL360232 Human DNA	941	18	2.3 166826	2	AL391707	AL391707 Human DNA
c 869	18	2.3 157931	2	AL627126	AL627126 Danio rer	942	18	2.3 166876	2	AL356864	AL356864 Homo sapi
c 870	18	2.3 158090	9	AC006353	AC006353 Homo sapi	c 943	18	2.3 166998	9	AC006230	AC006230 Homo sapi
871	18	2.3 158213	2	AF214635	AF214635 Homo sapi	c 944	18	2.3 167179	9	AC026182	AC026182 Homo sapi
c 872	18	2.3 158250	2	AC020781	AC020781 Homo sapi	945	18	2.3 167182	9	AC069208	AC069208 Homo sapi
873	18	2.3 158326	2	AC011333	AC011333 Homo sapi	c 946	18	2.3 167183	9	AC087086	AC087086 Homo sapi
874	18	2.3 158539	2	AC010622	AC010622 Homo sapi	947	18	2.3 167199	2	AC104430	AC104430 Homo sapi
c 875	18	2.3 158647	2	AC024478	AC024478 Homo sapi	948	18	2.3 167238	9	AC024903	AC024903 Homo sapi
c 876	18	2.3 158886	2	AC016944	AC016944 Homo sapi	949	18	2.3 167245	9	AC099521	AC099521 Homo sapi

[illegible]

ACCESSION AJ318878
VERSION AJ318878.1 GI:18157205
KEYWORDS XAGE-1c gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Zendman,A.J.W., van Kraats,A.A., Weidle,U.H., Ruiter,D.R. and Van
Muijen,G.N.P.
TITLE Expression profile of members of the XAGE cancer/testis antigen
family
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 740)
AUTHORS Zendman,A.J.W.
TITLE Direct Submission
JOURNAL Submitted (14-AUG-2001) Zendman A.J.W., Department of Pathology,
University Medical Center St Radboud, Geert Grooteplein Zuid 24, PO
Box 9101 6500 HB Nijmegen, NETHERLANDS
FEATURES
source
1..740
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp11.21-Xp11.22"
/tissue_type="Ewing sarcoma"
163..645
/gene="XAGE-1c"
163..645
/gene="XAGE-1c"
/codon_start=1
/product="XAGE-1c protein"
/protein_id="CAC82986.1"
/db_xref="GI:18157206"
/translation="MCHAHGPSCLVTATITREGGPRSGAQAIAKIGCCWGYSPRSTW
NPDRRTWPOTGEGRHERHTQNTASPRSPVESPKNQOLKLVGILHLGSRQK
KIRLQLRSQCAWKVICKSCISQTPGINDLGSQVVKIIPREHCKMPAGEEQPV
"
BASE COUNT 224 a 170 c 216 g 130 t
ORIGIN
Query Match 94.8%; Score 740; DB 9; Length 740;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 40 CTGGATTCTTTCCTCGTACTGAGACACGGCGGTAGGTCACAGGCAGATCCAACTGGG 99
Db 1 CTGGATTCTTTCCTCGTACTGAGACACGGCGGTAGGTCACAGGCAGATCCAACTGGG 60
Qy 100 AGTTGAAGTGTGAGTGTGAGAGTGAAGAGGAACACAGCGCTTCCGGAGGGTGTGTGTCTCA 159
Db 61 AGTTGAAGTGTGAGTGTGAGAGTGAAGAGGAACACAGCGCTTCCGGAGGGTGTGTGTCTCA 120
Qy 160 GTGACTCAGAGTGAAGAGGCCCTCGAAGTCGTCTGCTCCTCTCATGCGGTGCCAGGCCAT 219
Db 121 GTGACTCAGAGTGAAGAGGCCCTCGAAGTCGTCTGCTCCTCTCATGCGGTGCCAGGCCAT 180
Qy 220 GGACCTCTTCTGTCGTCACGGCCATAACTAGGAGGAAGAGGAGGCGCAGAGTGGAGGG 279
Db 181 GGACCTCTTCTGTCGTCACGGCCATAACTAGGAGGAAGAGGAGGCGCAGAGTGGAGGG 240
Qy 280 GCTCAGCGAAGCTGGGGTGTGTGTGGGGTATCCGAGTCCCAGAAGCACTGTGAACCCC 339
Db 241 GCTCAGCGAAGCTGGGGTGTGTGTGGGGTATCCGAGTCCCAGAAGCACTGTGAACCCC 300
Qy 340 GACAGAGATTCTGGACTCCCGACACGGGACCGAGGAGGAGCGGCGATGACGACACACA 399
Db 301 GACAGAGATTCTGGACTCCCGACACGGGACCGAGGAGGAGCGGCGATGACGACACACA 360
Qy 400 CAACACAGAACACACAGCGAGTCCCGAGGAGCCAGCTAATGGAGAGCCCAAAAGAG 459
Db 361 CAACACAGAACACACAGCGAGTCCCGAGGAGCCAGCTAATGGAGAGCCCAAAAGAG 420

Qy 460 AACGACGAGCTGAAAGTCGGGATCCTACACCTGGCGAGCAGACAGAAAGATCAGGATA 519
Db 421 AACGACGAGCTGAAAGTCGGGATCCTACACCTGGCGAGCAGACAGAAAGATCAGGATA 480
Qy 520 CAGCTGAGATCCCAGTGCAGACATGGAAGTGTATCTGCAAGAGCTGCATCAGTCAACA 579
Db 481 CAGCTGAGATCCCAGTGCAGACATGGAAGTGTATCTGCAAGAGCTGCATCAGTCAACA 540
Qy 580 CCGGGGATAAATCTGGATTTCGGGTTCCGGCGTCAAGGTGAAGATAATACCTAAAGAGAA 639
Db 541 CCGGGGATAAATCTGGATTTCGGGTTCCGGCGTCAAGGTGAAGATAATACCTAAAGAGAA 600
Qy 640 CACTGTAAAAATGCCAGAAGCAGGTGAAGAGCAACCAACCAAGTTTAAATGAAGACAAGCTGA 699
Db 601 CACTGTAAAAATGCCAGAAGCAGGTGAAGAGCAACCAACCAAGTTTAAATGAAGACAAGCTGA 660
Qy 700 AACACGCAAGCTGGTTTTATATTAGATATTGACTTAACTATCTCAATAAAGTTTTC 759
Db 661 AACACGCAAGCTGGTTTTATATTAGATATTGACTTAACTATCTCAATAAAGTTTTC 720
Qy 760 AGCTTTCCACCAAAAAA 779
Db 721 AGCTTTCCACCAAAAAA 740
RESULT 3
LOCUS BC009538 529 bp mRNA linear PRI 22-OCT-2001
DEFINITION Homo sapiens, clone MGC:10063 IMAGE:3893227, mRNA, complete cds.
ACCESSION BC009538
VERSION BC009538.1 GI:16306934
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 529)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (29-JUN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DFP/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAX Plate: 14 Row: g Column: 8.
FEATURES
source
1..529
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:10063 IMAGE:3893227"
/tissue_type="Lung, carcinoma, large cell
undifferentiated."
/clone_lib="NIH_MGC_69"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
184..429
/codon_start=1
/product="Unknown (protein for MGC:10063)"
/protein_id="AAH09538.1"

BASE COUNT		184 a	116 c	146 g	83 t
ORIGIN					
/db_xref="GI:16306935"					
/translation="MESPKNKKNQQLKVGILHLSRQKKIRIQLRSQCATWKVICKSCI					
SQTGGINLDLGSVKVKIIPKEEHCKMPEAGEEQPV"					
Query Match 67.3%; Score 526; DB 9; Length 529;					
Best Local Similarity 100.0%; Pred. No. 5.3e-233;					
Matches 526; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	256	GAAGGAGGGCCGAGGAGTGGAGGGCTCAGCGGAAGCTGGGTGCTGTGGGGCTATCCG	315		
Db	1	GAAGGAGGGCCGAGGAGTGGAGGGCTCAGCGGAAGCTGGGTGCTGTGGGGCTATCCG	60		
Qy	316	AGTCCCAGAAGCACTGGGAACCCCGACAGAAGATTCTGCAGTCCCAGACGGJACACGA	375		
Db	61	AGTCCCAGAAGCACTGGGAACCCCGACAGAAGATTCTGCAGTCCCAGACGGJACACGA	120		
Qy	376	GAGGAGCGGCATGAGCGACACACACAACACACAGCCAGTCCCGAGACGCCA	435		
Db	121	GAGGAGCGGCATGAGCGACACACACAACACACAGCCAGTCCCGAGACGCCA	180		
Qy	436	GTAATGAGAGCCCCCAAAAGAAGAACACAGCTGAAAGTCGGGATCCTACACCTGGGC	495		
Db	181	GTAATGAGAGCCCCCAAAAGAAGAACACAGCTGAAAGTCGGGATCCTACACCTGGGC	240		
Qy	496	AGCAGACAGAGAGATCAGGATACAGCTCAGATCCAGTCCGCGACATGGAAGGTGATC	555		
Db	241	AGCAGACAGAGAGATCAGGATACAGCTCAGATCCAGTCCGCGACATGGAAGGTGATC	300		
Qy	556	TGCAAGAGCTGCATCAGTCAACACCCGGGATAAATCTGGATTTGGGTTCCGCGCTCAAG	615		
Db	301	TGCAAGAGCTGCATCAGTCAACACCCGGGATAAATCTGGATTTGGGTTCCGCGCTCAAG	360		
Qy	616	GTGAAGATAATACCTAAGAGGACACTGTAAATGCCAGAAGCAGTCAAGAGCAACCA	675		
Db	361	GTGAAGATAATACCTAAGAGGACACTGTAAATGCCAGAAGCAGTCAAGAGCAACCA	420		
Qy	676	CAAGTTTAAATGAAGACAAGCTGAAACACCGGATTTTATATTAGATATTTGACT	735		
Db	421	CAAGTTTAAATGAAGACAAGCTGAAACACCGGATTTTATATTAGATATTTGACT	480		
Qy	736	TAAACTATCTCAATAAGATTTCCAGCTTTCACCAAAAAA	781		
Db	481	TAAACTATCTCAATAAGATTTTCAGCTTTCACCAAAAAA	526		
RESULT 4					
HSA400997					
LOCUS					
DEFINITION Homo sapiens XAGE-1b gene for cancer/testis-associated protein					
XAGE-1b, exons 1-4.					
ACCESSION AJ400997					
VERSION AJ400997.1 GI:13992557					
KEYWORDS cancer/testis-associated protein; XAGE-1b gene.					
SOURCE human.					
ORGANISM Homo sapiens					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE 1 (bases 1 to 6060)					
AUTHORS Zendenman,A.J.W., van Kraals,A.A., Weidie,U.H., Ruiter,D.J. and Van Muijen,G.N.P.					
TITLE XAGE-1b, a novel cancer/testis-associated gene, induced in melanoma metastases					
JOURNAL Unpublished					
REFERENCE 2 (bases 1 to 6060)					
AUTHORS Zendenman,A.J.W.					
TITLE Direct Submission					
JOURNAL Submitted (24-MAY-2000) Zendenman A.J.W., Pathology, University Medical Center St. Radboud, Geert Grooteplein Zuid 24, Nijmegen, 6500 HB POBox 9101, NETHERLANDS					
FEATURES					
source Location/Qualifiers					
1..6060					

exon		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/chromosome="X"
		/map="Xp11.21-Xp11.22"
		/tissue_type="melanoma metastasis"
		659..736
		/gene="XAGE-1b"
		/number=1
		659..5636
		/gene="XAGE-1b"
mRNA		Join(659..736,2143..2284,2812..2938,5529..5636)
intron		/gene="XAGE-1b"
		737..2142
		/gene="XAGE-1b"
		/number=1
exon		2143..2284
		/gene="XAGE-1b"
		/number=2
CDS		Join(2189..2284,2812..2938,5529..5551)
		/gene="XAGE-1b"
		/codon_start=1
		/product="cancer/testis-associated protein XAGE-1b"
		/protein_id="CAC38108.1"
		/db_xref="GI:13992558"
		/translation="MESPKNKKNQQLKVGILHLSRQKKIRIQLRSQCATWKVICKSCI
		SQTGGINLDLGSVKVKIIPKEEHCKMPEAGEEQPV"
intron		2285..2811
		/gene="XAGE-1b"
		/number=2
exon		2812..2938
		/gene="XAGE-1b"
		/number=3
intron		2939..5528
		/gene="XAGE-1b"
		/number=3
exon		5529..5636
		/gene="XAGE-1b"
		/number=4
polyA_signal		5614..5619
		/gene="XAGE-1b"
BASE COUNT		1700 a 1184 c 1420 g 1756 t
ORIGIN		
Query Match		50.2%; Score 392; DB 9; Length 6060;
Best Local Similarity		100.0%; Pred. No. 3.6e-171;
Matches 392; Conservative		0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	CGCGCGGAGCTGTGAGCGGGGACTCGCTCCCTGAGGTCTGGATTCTTTCTCCGCTACT 60
Db	665	CGCGCGGAGCTGTGAGCGGGGACTCGGTCCCTGAGGTCTGGATTCTTTCTCCGCTACT 724
Qy	61	GAGACACGGCGGTAGGTCCACAGGCAGATCCAACTGGGAGTTGAAGTGTGAGTGAAGT 120
Db	725	GAGACACGGCGGTAGGTCCACAGGCAGATCCAACTGGGAGTTGAAGTGTGAGTGAAGT 784
Qy	121	GAAGAGAACACAGCAGCTCCCGAGGGTGTCTGTCAGTCACTCAGAGTGAAGAGGCC 180
Db	785	GAAGAGAACACAGCAGCTCCCGAGGGTGTCTGTCAGTCACTCAGAGTGAAGAGGCC 844
Qy	181	CTCGAAGTGTGTCTCTCATTCGCGTCCACGGCCATGGACCTTCTTGTCTCGTCAAG 240
Db	845	CTCGAAGTGTGTCTCTCATTCGCGTCCACGGCCATGGACCTTCTTGTCTCGTCAAG 904
Qy	241	GCCATAACTAGGAGGAAGGGCCGAGAGTGGAGGGCTCAGCGGAAGCTGGGGTGC 300
Db	905	GCCATAACTAGGAGGAAGGGCCGAGAGTGGAGGGCTCAGCGGAAGCTGGGGTGC 964
Qy	301	TGTTGGGGGTATCCGAGTCCCAGAACACCTGGAAACCCCGACAGAGATTCTGGACTCCC 360
Db	965	TGTTGGGGGTATCCGAGTCCCAGAACACCTGGAAACCCCGACAGAGATTCTGGACTCCC 1024
Qy	361	CAGACGGGACACAGGAGGGACGGATGACG 392

Db	1025	CAGACGGGACGAGGAGCGGACCGCATGAGCG	1056
RESULT 5			
AC025553/C			
LOCUS	AC025553	180859 bp	DNA linear HTG 01-SEP-2000
DEFINITION	Homo sapiens chromosome X clone RP11-485B17, WORKING DRAFT		
SEQUENCE	SEQUENCE, 23 unordered pieces.		
ACCESSION	AC025553		
VERSION	AC025553.5	GI:9958270	
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 180859)		
JOURNAL	Waterston,R.H.		
AUTHORS	The sequence of Homo sapiens clone		
TITLE	Unpublished		
JOURNAL	2 (bases 1 to 180859)		
AUTHORS	Waterston,R.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (10-MAR-2000) Genome Sequencing Center, Washington		
AUTHORS	University School of Medicine, 4444 Forest Park Parkway, St. Louis,		
TITLE	MO 63108, USA		
COMMENT	On Sep 1, 2000 this sequence version replaced gi:9954853.		
	----- Genome Center -----		
	Center: Washington University Genome Sequencing Center		
	Center code: WUGSC		
	Web site:http://genome.wustl.edu/gsc/index.shtml		
	----- Project Information -----		
	Center project name: H.NH0485B17		
	----- Summary Statistics -----		
	Sequencing vector: M13; 100%		
	Chemistry: Dye-terminator; 0%		
	Chemistry: Dye-terminator Big Dye; 0% of reads		
	Assembly program: Phrap; version 0.990319		
	Consensus quality: 170599 bases at least Q40		
	Consensus quality: 173308 bases at least Q30		
	Consensus quality: 174949 bases at least Q20		
	Insert size: 187000; agarose-fp		
	Insert size: 178075; sum-of-contents		
	Quality coverage: 4.89 in Q20 bases; agarose-fp		
	Quality coverage: 5.22 in Q20 bases; sum-of-contents		

	* NOTE: This is a 'working draft' sequence. It currently		
	* consists of 23 contigs. The true order of the pieces		
	* is not known and their order in this sequence record is		
	* arbitrary. Gaps between the contigs are represented as		
	* runs of N, but the exact sizes of the gaps are unknown.		
	* This record will be updated with the finished sequence		
	* as soon as it is available and the accession number will		
	* be preserved.		

	1	1341:	contig of 1341 bp in length
	1342	1441:	gap of unknown length
	1442	2824:	contig of 1383 bp in length
	2825	2924:	gap of unknown length
	2925	4091:	contig of 1167 bp in length
	4092	4191:	gap of unknown length
	4192	5300:	contig of 1109 bp in length
	5301	5400:	gap of unknown length
	5401	6531:	contig of 1131 bp in length
	6532	6631:	gap of unknown length
	6632	7946:	contig of 1315 bp in length
	7947	8046:	gap of unknown length
	8047	9816:	contig of 1770 bp in length
	9817	9916:	gap of unknown length
	9917	11066:	contig of 1150 bp in length
	11067	11166:	gap of unknown length
	11167	12295:	contig of 1129 bp in length
	12296	12395:	gap of unknown length

	* NOTE: This is a 'working draft' sequence. It currently		
	* consists of 23 contigs. The true order of the pieces		
	* is not known and their order in this sequence record is		
	* arbitrary. Gaps between the contigs are represented as		
	* runs of N, but the exact sizes of the gaps are unknown.		
	* This record will be updated with the finished sequence		
	* as soon as it is available and the accession number will		
	* be preserved.		

	1	1341:	contig of 1341 bp in length
	1342	1441:	gap of unknown length
	1442	2824:	contig of 1383 bp in length
	2825	2924:	gap of unknown length
	2925	4091:	contig of 1167 bp in length
	4092	4191:	gap of unknown length
	4192	5300:	contig of 1109 bp in length
	5301	5400:	gap of unknown length
	5401	6531:	contig of 1131 bp in length
	6532	6631:	gap of unknown length
	6632	7946:	contig of 1315 bp in length
	7947	8046:	gap of unknown length
	8047	9816:	contig of 1770 bp in length
	9817	9916:	gap of unknown length
	9917	11066:	contig of 1150 bp in length
	11067	11166:	gap of unknown length
	11167	12295:	contig of 1129 bp in length
	12296	12395:	gap of unknown length

	* NOTE: This is a 'working draft' sequence. It currently		
	* consists of 23 contigs. The true order of the pieces		
	* is not known and their order in this sequence record is		
	* arbitrary. Gaps between the contigs are represented as		
	* runs of N, but the exact sizes of the gaps are unknown.		
	* This record will be updated with the finished sequence		
	* as soon as it is available and the accession number will		
	* be preserved.		

	1	1341:	contig of 1341 bp in length
	1342	1441:	gap of unknown length
	1442	2824:	contig of 1383 bp in length
	2825	2924:	gap of unknown length
	2925	4091:	contig of 1167 bp in length
	4092	4191:	gap of unknown length
	4192	5300:	contig of 1109 bp in length
	5301	5400:	gap of unknown length
	5401	6531:	contig of 1131 bp in length
	6532	6631:	gap of unknown length
	6632	7946:	contig of 1315 bp in length
	7947	8046:	gap of unknown length
	8047	9816:	contig of 1770 bp in length
	9817	9916:	gap of unknown length
	9917	11066:	contig of 1150 bp in length
	11067	11166:	gap of unknown length
	11167	12295:	contig of 1129 bp in length
	12296	12395:	gap of unknown length

	* NOTE: This is a 'working draft' sequence. It currently		
	* consists of 23 contigs. The true order of the pieces		
	* is not known and their order in this sequence record is		
	* arbitrary. Gaps between the contigs are represented as		
	* runs of N, but the exact sizes of the gaps are unknown.		
	* This record will be updated with the finished sequence		
	* as soon as it is available and the accession number will		
	* be preserved.		

	1	1341:	contig of 1341 bp in length
	1342	1441:	gap of unknown length
	1442	2824:	contig of 1383 bp in length
	2825	2924:	gap of unknown length
	2925	4091:	contig of 1167 bp in length
	4092	4191:	gap of unknown length
	4192	5300:	contig of 1109 bp in length
	5301	5400:	gap of unknown length
	5401	6531:	contig of 1131 bp in length
	6532	6631:	gap of unknown length
	6632	7946:	contig of 1315 bp in length
	7947	8046:	gap of unknown length
	8047	9816:	contig of 1770 bp in length
	9817	9916:	gap of unknown length
	9917	11066:	contig of 1150 bp in length
	11067	11166:	gap of unknown length
	11167	12295:	contig of 1129 bp in length
	12296	12395:	gap of unknown length

	* NOTE: This is a 'working draft' sequence. It currently		
	* consists of 23 contigs. The true order of the pieces		
	* is not known and their order in this sequence record is		
	* arbitrary. Gaps between the contigs are represented as		
	* runs of N, but the exact sizes of the gaps are unknown.		
	* This record will be updated with the finished sequence		
	* as soon as it is available and the accession number will		
	* be preserved.		

	1	1341:	contig of 1341 bp in length
	1342	1441:	gap of unknown length
	1442	2824:	contig of 1383 bp in length
	2825	2924:	gap of unknown length
	2925	4091:	contig of 1167 bp in length
	4092	4191:	gap of unknown length
	4192	5300:	contig of 1109 bp in length
	5301	5400:	gap of unknown length
	5401	6531:	contig of 1131 bp in length
	6532	6631:	gap of unknown length
	6632	7946:	contig of 1315 bp in length
	7947	8046:	gap of unknown length
	8047	9816:	contig of 1770 bp in length
	9817	9916:	gap of unknown length
	9917	11066:	contig of 1150 bp in length
	11067	11166:	gap of unknown length
	11167	12295:	contig of 1129 bp in length
	12296	12395:	gap of unknown length

	* NOTE: This is a 'working draft' sequence. It currently		
	* consists of 23 contigs. The true order of the pieces		
	* is not known and their order in this sequence record is		
	* arbitrary. Gaps between the contigs are represented as		
	* runs of N, but the exact sizes of the gaps are unknown.		
	* This record will be updated with the finished sequence		
	* as soon as it is available and the accession number will		
	* be preserved.		

	1	1341:	contig of 1341 bp in length
	1342	1441:	gap of unknown length
	1442	2824:	contig of 1383 bp in length
	2825	2924:	gap of unknown length
	2925	4091:	contig of 1167 bp in length
	4092	4191:	gap of unknown length
	4192	5300:	contig of 1109 bp in length
	5301	5400:	gap of unknown length
	5401	6531:	contig of 1131 bp in length
	6532	6631:	gap of unknown length
	6632	7946:	contig of 1315 bp in length
	7947	8046:	gap of unknown length
	8047	9816:	contig of 1770 bp in length
	9817	9916:	gap of unknown length
	9917	11066:	contig of 1150 bp in length
	11067	11166:	gap of unknown length
	11167	12295:	contig of 1129 bp in length
	12296	12395:	gap of unknown length

	* NOTE: This is a 'working draft' sequence. It currently		
	* consists of 23 contigs. The true order of the pieces		
	* is not known and their order in this sequence record is		
	* arbitrary. Gaps between the contigs are represented as		
	* runs of N, but the exact sizes of the gaps are unknown.		
	* This record will be updated with the finished sequence		
	* as soon as it is available and the accession number will		
	* be preserved.		

	1	1341:	contig of 1341 bp in length
	1342	1441:	gap of unknown length
	1442	2824:	contig of 1383 bp in length
	2825	2924:	gap of unknown length
	2925	4091:	contig of 1167 bp in length
	4092	4191:	gap of unknown length
	4192	5300:	contig of 1109 bp in length
	5301	5400:	gap of unknown length
	5401	6531:	contig of 1131 bp in length
	6532	6631:	gap of unknown length
	6632	7946:	contig of 1315 bp in length
	7947	8046:	gap of unknown length
	8047	9816:	contig of 1770 bp in length
	9817	9916:	gap of unknown length
	9917	11066:	contig of 1150 bp in length
	11067	11166:	gap of unknown length
	11167	12295:	contig of 1129 bp in length
	12296	12395:	gap of unknown length

	* NOTE: This is a 'working draft' sequence. It currently		
	* consists of 23 contigs. The true order of the pieces		
	* is not known and their order in this sequence record is		
	* arbitrary. Gaps between the contigs are represented as		
	* runs of N, but the exact sizes of the gaps are unknown.		
	* This record will be updated with the finished sequence		
	* as soon as it is available and the accession number will		
	* be preserved.		

	1	1341:	contig of 1341 bp in length
	1342	1441:	gap of unknown length
	1442	2824:	contig of 1383 bp in length
	2825	2924:	gap of unknown length
	2925	4091:	contig of 1167 bp in length
	4092	4191:	gap of unknown length
	4192	5300:	contig of 1109 bp in length
	5301	5400:	gap of unknown length
	5401	6531:	contig of 1131 bp in length
	6532	6631:	gap of unknown length
	6632	7946:	contig of 1315 bp in length
	7947	8046:	gap of unknown length
	8047	9816:	contig of 1770 bp in length
	9817	9916:	gap of unknown length
	9917	11066:	contig of 1150 bp in length
	11067	11166:	gap of unknown length
	11167	12295:	contig of 1129 bp in length
	12296	12395:	gap of unknown length

	* NOTE: This is a 'working draft' sequence. It currently		
	* consists of 23 contigs. The true order of the pieces		
	* is not known and their order in this sequence record is		
	* arbitrary. Gaps between the contigs are represented as		
	* runs of N, but the exact sizes of the gaps are unknown.		
	* This record will be updated with the finished sequence		
	* as soon as it is available and the accession number will		
	* be preserved.		

	1	1341:	contig of 1341 bp in length
	1342	1441:	gap of unknown length
	1442	2824:	contig of 1383 bp in length
	2825	2924:	gap of unknown length
	2925	4091:	contig of 1167 bp in length
	4092	4191:	gap of unknown length
	4192	5300:	contig of 1109 bp in length
	5301	5400:	gap of unknown length
	5401	6531:	contig of 1131 bp in length
	6532	6631:	gap of unknown length
	6632	7946:	contig of 1315 bp in length
	7947	8046:	gap of unknown length
	8047	9816:	contig of 1770 bp in length
	9817	9916:	gap of unknown length
	9917	11066:	contig of 1150 bp in length
	11067	11166:	gap of unknown length
	11167	12295:	contig of 1129 bp in length
	12296	12395:	gap of unknown length

	* NOTE: This is a 'working draft' sequence. It currently		
	* consists of 23 contigs. The true order of the pieces		
	* is not known and their order in this sequence record is		
	* arbitrary. Gaps between the contigs are represented as		
	* runs of N, but the exact sizes of the gaps are unknown.		
	* This record will be updated with the finished sequence		
	* as soon as it is available and the accession number will		
	* be preserved.		

	1	1341:	contig of 1341 bp in length
	1342	1441:	gap of unknown length
	1442	2824:	contig of 1383 bp in length
	2825	2924:	gap of unknown length
	2925	4091:	contig of 1167 bp in length
	4092	4191:	gap of unknown length
	4192	5300:	contig of 1109 bp in length
	5301	5400:	gap of unknown length
	5401	6531:	contig of 1131 bp in length
	6532	6631:	gap of unknown length
	6632	7946:	contig of 1315 bp in length
	7947	8046:	gap of unknown length
	8047	9816:	contig of 1770 bp in length
	9817	9916:	gap of unknown length
	9917	11066:	contig of 1150 bp in length
	11067	11166:	gap of unknown length
	11167	12295:	contig of 1129 bp in length
	12296	12395:	gap of unknown length

	* NOTE: This is a 'working draft' sequence. It currently		
	* consists of 23 contigs. The true order of the pieces		
	* is not known and their order in this sequence record is		
	* arbitrary. Gaps between the contigs are represented as		
	* runs of N, but the exact sizes of the gaps are unknown.		
	* This record will be updated with the finished sequence		
	* as soon as it is available and the accession number will		
	* be preserved.		

	1	1341:	contig of 1341 bp in length
	1342	1441:	gap of unknown length
	1442	2824:	contig of 1383 bp in length
	2825	2924:	gap of unknown length
	2925	4091:	contig of 1167 bp in length
	4092	4191:	gap of unknown length
	4192	5300:	contig of 1109 bp in length
	5301	5400:	gap of unknown length
	5401	6531:	contig of 1131 bp in length
	6532	6631:	gap of unknown length
	6632	7946:	contig of 1315 bp in length
	7947	8046:	gap of unknown length
	8047	9816:	contig of 1770 bp in length
	9817	9916:	gap of unknown length
	9917	11066:	contig of 1150 bp in length
	11067	11166:	gap of unknown length
	11167	12295:	contig of 1129 bp in length
	12296	12395:	gap of unknown length

	* NOTE: This is a 'working draft' sequence. It currently		
	* consists of 23 contigs. The true order of the pieces		
	* is not known and their order in this sequence record is		
	* arbitrary. Gaps between the contigs are represented as		
	* runs of N, but the exact sizes of the gaps are unknown.		
	* This record will be updated with the finished sequence		
	* as soon as it is available and the accession number will		
	* be preserved.		

	1	1341:	contig of 1341 bp in length
	1342	1441:	gap of unknown length
	1442	2824:	contig of 1383 bp in length
	2825	2924:	gap of unknown length
	2925	4091:	contig of 1167 bp in length
	4092	4191:	gap of unknown length
	4192	5300:	contig of 1109 bp in length
	5301	5400:	gap of unknown length
	5401	6531:	contig of 1131 bp in length
	6532	6631:	gap of unknown length

```
misc_feature 50028..69100
/note="assembly_name:Contig58"
misc_feature 69201..90260
/note="assembly_name:Contig59"
misc_feature 90361..141786
/note="assembly_name:Contig60"
misc_feature 141887..180859
/note="assembly_name:Contig61"
BASE COUNT 53372 a 37170 c 37454 g 50639 t 2224 others
ORIGIN
Query Match 50.2%; Score 392; DB 2; Length 180859;
Best Local Similarity 100.0%; Pred. No. 1.4e-171;
Matches 392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCGGAGCTGTGAGCCGGGACTCGGGTCCCTGAGTCTGGATTCTTTCTCCGCTACT 60
Db 157680 GCGCGGAGCTGTGAGCCGGGACTCGGGTCCCTGAGTCTGGATTCTTTCTCCGCTACT 157621

QY 61 GAGACACGCGGGTAGGTCCACAGCAGATCAACTGGGAGTTGAAGTGTGAGTGAGAGT 120
Db 157620 GAGACACGCGGGTAGGTCCACAGCAGATCAACTGGGAGTTGAAGTGTGAGTGAGAGT 157561

QY 121 GAAGAGAACACAGCAGGCTTCGGAGGTTGTGTGTCAGTCACTGACAGTGAAGAGCC 180
Db 157560 GAAGAGAACACAGCAGGCTTCGGAGGTTGTGTGTCAGTCACTGACAGTGAAGAGCC 157501

QY 181 CTCGAAGTCGTCCCTCTCATGCGGTGCCACGCCATGACCTTCTTGTCTCGTCACG 240
Db 157500 CTCGAAGTCGTCCCTCTCATGCGGTGCCACGCCATGACCTTCTTGTCTCGTCACG 157441

QY 241 GCATTAACTAGGAGAGGAGGCGGAGGAGTGTGTCAGTCACTGACAGTGAAGAGCC 300
Db 157440 GCATTAACTAGGAGAGGAGGCGGAGGAGTGTGTCAGTCACTGACAGTGAAGAGCC 157381

QY 301 TGTGGGGGTATCCGAGTCCCAAGACACCTGGAACCCGACAGAGATTTCTGGACTCCC 360
Db 157380 TGTGGGGGTATCCGAGTCCCAAGACACCTGGAACCCGACAGAGATTTCTGGACTCCC 157321

QY 361 CACAGCGGACAGGAGAGGAGCGGATGAGCG 392
Db 157320 CACAGCGGACAGGAGAGGAGCGGATGAGCG 157289

RESULT 6
AX245237
LOCUS AX245237 396 bp DNA linear PAT 28-SEP-2001
DEFINITION Sequence 167 from Patent WO0166753.
ACCESSION AX245237
VERSION AX245237.1 GI:15859911
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 396)
AUTHORS Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D.,
Sudduth-Klinger,J., Reinhard,C., Randazzo,F., Kennedy,G.C., Pot,D.,
Kassam,A., Lamson,G., Drmanac,R., Crkvenjakov,R., Dickson,M.,
Drmanac,S., Labat,I., Leshkowitz,D., Kita,D., Garcia,V. and
Stache-Crain,B.
TITLE Human genes and gene expression products
JOURNAL Patent: WO 0166753-A 167 13-SEP-2001.
Chiron Corporation (US) ; Hyseq Inc. (US)
FEATURES
source
1..396
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 82 a 97 c 146 g 71 t
ORIGIN
Query Match 50.1%; Score 391; DB 6; Length 396;
Best Local Similarity 100.0%; Pred. No. 2.4e-170;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

misc_feature 50028..69100
/note="assembly_name:Contig58"
misc_feature 69201..90260
/note="assembly_name:Contig59"
misc_feature 90361..141786
/note="assembly_name:Contig60"
misc_feature 141887..180859
/note="assembly_name:Contig61"
BASE COUNT 53372 a 37170 c 37454 g 50639 t 2224 others
ORIGIN
Query Match 49.9%; Score 390; DB 9; Length 626;
Best Local Similarity 100.0%; Pred. No. 6e-170;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCGCGGAGCTGTGAGCCGGGACTCGGGTCCCTGAGTCTGGATTCTTTCTCCGCTACTG 61
Db 1 CCGCGGAGCTGTGAGCCGGGACTCGGGTCCCTGAGTCTGGATTCTTTCTCCGCTACTG 60

QY 62 AGACACGCGGGTAGGTCCACAGCAGATCAACTGGGAGTTGAAGTGTGAGTGAGAGTG 121
Db 61 AGACACGCGGGTAGGTCCACAGCAGATCAACTGGGAGTTGAAGTGTGAGTGAGAGTG 120

QY 122 AAGAGGAACACAGCAGGCTTCGGAGGTTGTGTGTCAGTCACTGACAGTGAAGAGCC 181
Db 121 AAGAGGAACACAGCAGGCTTCGGAGGTTGTGTGTCAGTCACTGACAGTGAAGAGCC 180

QY 182 TCGAAGTCGTCCCTCTCATGCGGTGCCACGCCATGACCTTCTTGTCTCGTCACGG 241
Db 181 TCGAAGTCGTCCCTCTCATGCGGTGCCACGCCATGACCTTCTTGTCTCGTCACGG 240

QY 242 CCATAACTAGGAGAGGAGGCGGAGGAGTGTGTCAGTCACTGACAGTGAAGAGCC 301
Db 241 CCATAACTAGGAGAGGAGGCGGAGGAGTGTGTCAGTCACTGACAGTGAAGAGCC 300

QY 302 GTTGGGGGTATCCGAGTCCCAAGACACCTGGAACCCGACAGAGATTTCTGGACTCCC 361
Db 301 GTTGGGGGTATCCGAGTCCCAAGACACCTGGAACCCGACAGAGATTTCTGGACTCCC 360

QY 362 AGACGGGACAGGAGAGGAGCGGATGAGCG 392
Db 361 AGACGGGACAGGAGAGGAGCGGATGAGCG 391

RESULT 7
AF251237
LOCUS AF251237 626 bp mRNA linear PRI 23-AUG-2000
DEFINITION Homo sapiens XAGE-1 mRNA, complete cds.
ACCESSION AF251237
VERSION AF251237.1 GI:9885324
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 626)
AUTHORS Liu,X.F., Helman,L.J., Yeung,C., Lee,B. and Pastan,I.
XAGE-1, A New Gene That is Frequently Expressed in Ewing's Sarcoma
JOURNAL Unpublished
2 (bases 1 to 626)
AUTHORS Liu,X.F., Helman,L.J., Yeung,C., Lee,B. and Pastan,I.
Direct Submission
JOURNAL Submitted (31-MAR-2000) Lab of Mol. Biol., NCI, NIH, 37 Convent Dr.
MSC 4255, Bethesda, MD 20892, USA
FEATURES
source
1..626
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/tissue_type="testis"
86..526
/note="expressed in Ewing's sarcoma"
/codon_start=1
/product="XAGE-1"
/protein_id="AAG01401.1"
/db_xref="GI:9885325"
/translation="MLLKPQCACSLGVFSPAPSVYVTRRSCPEPATRVPEVWILSP
LLRHGHHTQNTHTASPRSPVMSPKKKNQKLVGILHLSRKKIRIQLRSOCATWK
VICKSCISQTPGINLDLGSVGVKVIIPREHCKMPEAGEEQPV"
BASE COUNT 194 a 153 c 170 g 109 t
ORIGIN
Query Match 49.9%; Score 390; DB 9; Length 626;
Best Local Similarity 100.0%; Pred. No. 6e-170;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 392 GACACACAAACACAGAACACACAGCCAGTCCCGAGGAGCCAGTAATGGAGAGCCCA 451
|||||
Db 234 GACACACAAACACAGAACACACAGCCAGTCCCGAGGAGCCAGTAATGGAGAGCCCA 293
|||||
QY 452 AAAAGAGAACACAGAGCTGAAAGTCGGGATCCTACACCTGGCGACGACACAGAGAAGA 511
|||||
Db 294 AAAAGAGAACACAGAGCTGAAAGTCGGGATCCTACACCTGGCGACGACACAGAGAAGA 353
|||||
QY 512 TCAGGATACAGCTGAGATCCCGAGTGGCGACATGGAAGTGATCTGCAAGAGCTGCATCA 571
|||||
Db 354 TCAGGATACAGCTGAGATCCCGAGTGGCGACATGGAAGTGATCTGCAAGAGCTGCATCA 413
|||||
QY 572 GTCAACACCCGGGATAAATCTGGATTTGGTTCCGGCGTCAAGTGGAAGTAATACCTTA 631
|||||
Db 414 GTCAACACCCGGGATAAATCTGGATTTGGTTCCGGCGTCAAGTGGAAGTAATACCTTA 473
|||||
QY 632 AAGAGGAACACTGTAATATGCCAGAGCAGGTGAAGAGCAACACAAAGTTTAAATGAAGA 691
|||||
Db 474 AAGAGGAACACTGTAATATGCCAGAGCAGGTGAAGAGCAACACAAAGTTTAAATGAAGA 533
|||||
QY 692 CAAGCTGAACACGCAAGCTGGTTTATATTAGATATTGGACTTAAACTATCTCAATAA 751
|||||
Db 534 CAAGCTGAACACGCAAGCTGGTTTATATTAGATATTGGACTTAAACTATCTCAATAA 593
|||||
QY 752 AGTTTTCAGCTTTCACCAAAAAA 781
|||||
Db 594 AGTTTTCAGCTTTCACCAAAAAA 623
|||||
RESULT 8
HSA290447 463 bp mRNA linear PRI 05-MAY-2001
LOCUS Homo sapiens mRNA for 9 kD cancer/testis associated protein
DEFINITION (Xage-lb).
ACCESSION AJ290447
VERSION AJ290447.1 GI:13992498
KEYWORDS cancer/testis associated protein; CTP9 gene.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 463)
Zendman,A.J.W., van Kraats,A.A., Weidle,U.H., Ruiter,D.J. and Van
Muijen,G.N.P.
CTP9, a novel cancer/testis-associated gene, induced in melanoma
metastases
TITLE Unpublished
JOURNAL 2 (bases 1 to 463)
REFERENCE Zendman,A.J.W.
AUTHORS Direct Submission
TITLE Submitted (04-MAY-2000) Zendman A.J.W., Pathology, University
JOURNAL Medical Center St. Radboud, Geert Grooteplein Zuid 24, Nijmegen,
6500 HB POBox 9101, NETHERLANDS
REMARK Revised by author 22-JAN-2001
FEATURES Location/Qualifiers
source 1..463
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp11.21-Xp11.22"
/tissue_type="melanoma metastasis"
125..370
/gene="Xage-lb"
125..370
/gene="Xage-lb"
/codon_start=1
/product="9 kD cancer/testis associated protein"
/protein_id="CAC38107.1"
/db_xref="GI:13992498"
/translation="MESPKKKNQOLKVGLHLGSRQKKRIQLRSOCATWKVICKSCI
SQTPGINLDGSGVKVRIIPKEEHCKMPAGEEQPV"
BASE COUNT 155 a 106 c 120 g 82 t
ORIGIN

Query Match 49.4%; Score 386; DB 9; Length 463;
Best Local Similarity 100.0%; Pred. No. 4.7e-168; Indels 0; Gaps 0;
Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 392 GACACACAAACACAGAACACACAGCCAGTCCCGAGGAGCCAGTAATGGAGAGCCCA 451
|||||
Db 78 GACACACAAACACAGAACACACAGCCAGTCCCGAGGAGCCAGTAATGGAGAGCCCA 137
|||||
QY 452 AAAAGAGAACACAGAGCTGAAAGTCGGGATCCTACACCTGGCGACGACACAGAGAAGA 511
|||||
Db 138 AAAAGAGAACACAGAGCTGAAAGTCGGGATCCTACACCTGGCGACGACAGAGAAGA 197
|||||
QY 512 TCAGGATACAGCTGAGATCCCGAGTGGCGACATGGAAGTGATCTGCAAGAGCTGCATCA 571
|||||
Db 198 TCAGGATACAGCTGAGATCCCGAGTGGCGACATGGAAGTGATCTGCAAGAGCTGCATCA 257
|||||
QY 572 GTCAACACCCGGGATAAATCTGGATTTGGTTCCGGCGTCAAGTGGAAGTAATACCTTA 631
|||||
Db 258 GTCAACACCCGGGATAAATCTGGATTTGGTTCCGGCGTCAAGTGGAAGTAATACCTTA 317
|||||
QY 632 AAGAGGAACACTGTAATATGCCAGAGCAGGTGAAGAGCAACACAAAGTTTAAATGAAGA 691
|||||
Db 318 AAGAGGAACACTGTAATATGCCAGAGCAGGTGAAGAGCAACACAAAGTTTAAATGAAGA 377
|||||
QY 692 CAAGCTGAACACGCAAGCTGGTTTATATTAGATATTGGACTTAAACTATCTCAATAA 751
|||||
Db 378 CAAGCTGAACACGCAAGCTGGTTTATATTAGATATTGGACTTAAACTATCTCAATAA 437
|||||
QY 752 AGTTTTCAGCTTTCACCAAAAAA 777
|||||
Db 438 AGTTTTCAGCTTTCACCAAAAAA 463
|||||
RESULT 9
AX062442 399 bp DNA linear PAT 24-JAN-2001
LOCUS Sequence 69 from Patent WO0100828.
DEFINITION AX062442
ACCESSION AX062442
VERSION AX062442.1 GI:12540317
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 399)
AUTHORS Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S.,
Carter,D., Retter,M.W. and Mannion,J.
Compositions and methods for the therapy and diagnosis of lung
cancer
TITLE Patent: WO 0100828-A 69 04-JAN-2001;
JOURNAL CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1..399
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 150 a 87 c 94 g 67 t 1 others
ORIGIN
Query Match 48.7%; Score 380; DB 6; Length 399;
Best Local Similarity 100.0%; Pred. No. 3e-165;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 392 GACACACAAACACAGAACACACAGCCAGTCCCGAGGAGCCAGTAATGGAGAGCCCA 451
|||||
Db 10 GACACACAAACACAGAACACACAGCCAGTCCCGAGGAGCCAGTAATGGAGAGCCCA 69
|||||
QY 452 AAAAGAGAACACAGAGCTGAAAGTCGGGATCCTACACCTGGCGACGACACAGAGAAGA 511
|||||
Db 70 AAAAGAGAACACAGAGCTGAAAGTCGGGATCCTACACCTGGCGACGACACAGAGAAGA 129
|||||
QY 512 TCAGGATACAGCTGAGATCCCGAGTGGCGACATGGAAGTGATCTGCAAGAGCTGCATCA 571
|||||

Db	130	TCAGGTACAGCTGAGATCCAGTCCAGTGCAGACATGGAAGGTGATCTGCAAGAGCTGCATCA	189
Qy	572	GTCAACACCGGGGATAAATCTGATTTGGTTCCGGCGTCAAGGTCAAGTAAATACCTA	631
Db	190	GTCAACACCGGGGATAAATCTGATTTGGTTCCGGCGTCAAGGTCAAGTAAATACCTA	249
Qy	632	AAGAGGAACACTGTAAATGCCAGACAGCTGGAAGAGCAACCAACAGTTTAAATGAAGA	691
Db	250	AAGAGGAACACTGTAAATGCCAGACAGCTGGAAGAGCAACCAACAGTTTAAATGAAGA	309
Qy	692	CAAGCTGAACACCGCAAGCTGTTTTATATTAGATATTTGACTTAACTATCTCAATAA	751
Db	310	CAAGCTGAACACCGCAAGCTGTTTTATATTAGATATTTGACTTAACTATCTCAATAA	369
Qy	752	AGTTTTCAGCTTTTCAACAA	771
Db	370	AGTTTTCAGCTTTTCAACAA	389
RESULT 10			
AX063163			
LOCUS	AX063163	Sequence 790 from Patent WO0100828.	457 bp DNA linear PAT 24-JAN-2001
DEFINITION	AX063163		
ACCESSION	AX063163		
VERSION	AX063163.1	GI:12541017	
KEYWORDS			
SOURCE		human.	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
TITLE		Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S.,	
JOURNAL		Carter, D., Retter, M.W. and Mannion, J.	
FEATURES		Compositions and methods for the therapy and diagnosis of lung	
source		cancer	
PATENT		Patent: WO 0100828-A 790 04-JAN-2001;	
CORIXA		CORIXA CORPORATION (US)	
Location/Qualifiers			
1.	457		
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
BASE COUNT	146 a	105 c	121 g 85 t
ORIGIN			
Query Match	47.1%;	Score 368;	DB 6; Length 457;
Best Local Similarity	100.0%;	Pred. No. 1e-159;	
Matches	368;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
Qy	392	GACACACAAACACAGAACACACAGCCAGTCCCGAGGAGCCAGTAATGGAGAGCCCA	451
Db	90	GACACACAAACACAGAACACACAGCCAGTCCCGAGGAGCCAGTAATGGAGAGCCCA	149
Qy	452	AAAGAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTGGCAGCAGACAGAGAAGA	511
Db	150	AAAGAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTGGCAGCAGACAGAGAAGA	209
Qy	512	TCAGGATACAGCTGAGATCCCGATCCCGCAGCATGGAAGGTGATCTGCAAGAGCTGCATCA	571
Db	210	TCAGGATACAGCTGAGATCCCGATCCCGCAGCATGGAAGGTGATCTGCAAGAGCTGCATCA	269
Qy	572	GTCAACACCGGGGATAAATCTGGATTTCCGGCTCAAGTGAAGATAATACCTA	631
Db	270	GTCAACACCGGGGATAAATCTGGATTTCCGGCTCAAGTGAAGATAATACCTA	329
Qy	632	AAGAGGAACACTGTAAATGCCAGACAGCTGGAAGAGCAACCAACAGTTTAAATGAAGA	691
Db	330	AAGAGGAACACTGTAAATGCCAGACAGCTGGAAGAGCAACCAACAGTTTAAATGAAGA	389
Qy	692	CAAGCTGAACACCGCAAGCTGTTTTATATTAGATATTTGACTTAACTATCTCAATAA	751
Db	390	CAAGCTGAACACCGCAAGCTGTTTTATATTAGATATTTGACTTAACTATCTCAATAA	449
Qy	752	AGTTTTCG 759	

Db	450	AGTTTTCG 457	
RESULT 11			
HS318879			
LOCUS	HS318879	Homo sapiens mRNA for XAGE-1d protein.	481 bp mRNA linear PRI 14-JAN-2002
DEFINITION	HS318879		
ACCESSION	AJ318879		
VERSION	AJ318879.1	GI:18157207	
KEYWORDS		XAGE-1d gene.	
SOURCE		human.	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
TITLE		Zendman, A.J.W., van Kraats, A.A., Weidle, U.H., Ruiter, D.R. and Van	
JOURNAL		Muijen, G.N.P.	
REFERENCE		Expression profile of members of the XAGE cancer/testis antigen	
AUTHORS		family	
TITLE		Unpublished	
JOURNAL		2 (bases 1 to 481)	
DEFINITION		Zendman, A.J.W.	
ACCESSION		Direct Submission	
VERSION		Submitted (14-AUG-2001) Zendman A.J.W., Department of Pathology,	
KEYWORDS		University Medical Center St Radboud, Geert Grooteplein Zuid 24, PO	
SOURCE		Box 9101 6500 HB Nijmegen, NETHERLANDS	
FEATURES		Location/Qualifiers	
1.	481		
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/chromosome="X"			
/map="Xp11.21-Xp11.22"			
/tissue="Ewing sarcoma"			
125..334			
/gene="XAGE-1d"			
125..334			
/gene="XAGE-1d"			
/codon_start=1			
/product="XAGE-1d protein"			
/protein_id="CAC82987.1"			
/db_xref="GI:18157208"			
/translation="MESPKKKNQQLKVGILHLSRQKKIRIQLRQVLCREMRDMEGD			
LOELHOSNTGDKSGFGFRQGEDNT"			
BASE COUNT	163 a	107 c	127 g 84 t
ORIGIN			
Query Match	31.4%;	Score 245;	DB 9; Length 481;
Best Local Similarity	100.0%;	Pred. No. 1.1e-102;	
Matches	245;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
Qy	535	TGCGCAGATGGAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACGGGGATAATCTG	594
Db	237	TGCGCAGATGGAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACGGGGATAATCTG	296
Qy	595	GATTTGGTTCCCGCGTCAAGGTGAAGATAATACCTTAAGAGCAACAGTGAAGTGGG	654
Db	297	GATTTGGTTCCCGCGTCAAGGTGAAGATAATACCTTAAGAGCAACAGTGAAGTGGG	356
Qy	655	GAACAGGTGAAGAGCAACCAAGTTTAAATGAAGCAACAGCTGAAGTGAAGTGGG	714
Db	357	GAACAGGTGAAGAGCAACCAAGTTTAAATGAAGCAACAGTGAAGTGAAGTGGG	416
Qy	715	TTTATATTAGATATTTGACTTAACTATCTCAATAAAGTTTTCACCAAAA	774
Db	417	TTTATATTAGATATTTGACTTAACTATCTCAATAAAGTTTTCACCAAAA	476
Qy	775	AAAAA 779	
Db	477	AAAAA 481	
RESULT 12			

```

AX063165
LOCUS AX063165 461 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 792 from Patent WO0100828.
ACCESSION AX063165
VERSION AX063165.1 GI:12541018
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 461)
AUTHORS Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S.,
Carter,D., Retter,M.W. and Mannion,J.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: WO 0100828-A 792 04-JAN-2001;
CORIXA CORPORATION (US)
FEATURES
source
1. .461
Location/Qualifiers
BASE COUNT 150 a 104 c 123 g 84 t
ORIGIN
Query Match 29.7%; Score 232; DB 6; Length 461;
Best Local Similarity 100.0%; Pred. No. 1.2e-96;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 535 TCGCGACATGGAAGGTGATCTGCAAGAGCTGTCATCAGTCAACACCGGGGATAAAATCTG 594
Db 230 TCGCGACATGGAAGGTGATCTGCAAGAGCTGTCATCAGTCAACACCGGGGATAAAATCTG 289

Qy 595 GATTGGGTTCCGGCGTCAAGGTGAAGATAATACCTAAAGAGGACACTGTAAATGCCA 654
Db 290 GATTGGGTTCCGGCGTCAAGGTGAAGATAATACCTAAAGAGGACACTGTAAATGCCA 349

Qy 655 GAAGCAGGTGAAGAGCAACACCAAGTTTAAATGAAGCAAGCTGAAACACGCAAGCTGG 714
Db 350 GAAGCAGGTGAAGAGCAACCAAGTTTAAATGAAGCAAGCTGAAACACGCAAGCTGG 409

Qy 715 TTTTATATTAGATTGACTTAACTATCTCAATAAAGTTTTCAGCTTTTC 766
Db 410 TTTTATATTAGATTGACTTAACTATCTCAATAAAGTTTTCAGCTTTTC 461

RESULT 13
AX062443
LOCUS AX062443 479 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 70 from Patent WO0100828.
ACCESSION AX062443
VERSION AX062443.1 GI:12540318
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 479)
AUTHORS Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S.,
Carter,D., Retter,M.W. and Mannion,J.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: WO 0100828-A 70 04-JAN-2001;
CORIXA CORPORATION (US)
FEATURES
source
1. .479
Location/Qualifiers
BASE COUNT 163 a 107 c 125 g 84 t
ORIGIN
Query Match 24.5%; Score 191; DB 6; Length 479;
Best Local Similarity 100.0%; Pred. No. 1.3e-77;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 535 TCGCGACATGGAAGGTGATCTGCAAGAGCTGTCATCAGTCAACACCGGGGATAAAATCTG 594
Db 232 TCGCGACATGGAAGGTGATCTGCAAGAGCTGTCATCAGTCAACACCGGGGATAAAATCTG 291
Qy 595 GATTGGGTTCCGGCGTCAAGGTGAAGATAATACCTAAAGAGGACACTGTAAATGCCA 654
Db 292 GATTGGGTTCCGGCGTCAAGGTGAAGATAATACCTAAAGAGGACACTGTAAATGCCA 351
Qy 655 GAAGCAGGTCAAGAGCAACACCAAGTTTAAATGAAGCAAGCTGAAACACGCAAGCTGG 714
Db 352 GAAGCAGGTCAAGAGCAACCAAGTTTAAATGAAGCAAGCTGAAACACGCAAGCTGG 411
Qy 715 TTTTATATTAG 725
Db 412 TTTTATATTAG 422

RESULT 14
AC016835
LOCUS AC016835 81953 bp DNA linear HTG 13-JUL-2000
DEFINITION Homo sapiens clone RP11-1P5, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC016835
VERSION AC016835.2 GI:9119883
KEYWORDS HTG; HTGS_PHASE0.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 81953)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens, clone RP11-1P5
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 81953)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Bozulavkiy,L., Boukhgalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArelano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferrel,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Gage,M.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heath,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Tirrell,A., Vassiliev,H., Vo.A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (08-DEC-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6539374.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RW/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2520
Center clone name: L_P_5
-----
* NOTE: This record contains 91 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that

```


* the record is updated, the accession number will
* be preserved.
* 1 769: contig of 769 bp in length
* 770 869: gap of 100 bp
* 870 1691: contig of 822 bp in length
* 1692 1791: gap of 100 bp
* 1792 2590: contig of 799 bp in length
* 2591 2690: gap of 100 bp
* 2691 3502: contig of 812 bp in length
* 3503 3602: gap of 100 bp
* 3603 4403: contig of 801 bp in length
* 4404 4503: gap of 100 bp
* 4504 5271: contig of 768 bp in length
* 5272 5371: gap of 100 bp
* 5372 6160: contig of 789 bp in length
* 6161 6260: gap of 100 bp
* 6261 7048: contig of 788 bp in length
* 7049 7148: gap of 100 bp
* 7149 7945: contig of 797 bp in length
* 7946 8045: gap of 100 bp
* 8046 8864: contig of 819 bp in length
* 8865 8964: gap of 100 bp
* 8965 9768: contig of 804 bp in length
* 9769 9868: gap of 100 bp
* 9869 10679: contig of 811 bp in length
* 10680 10779: gap of 100 bp
* 10780 11545: contig of 766 bp in length
* 11546 11645: gap of 100 bp
* 11646 12460: contig of 815 bp in length
* 12461 12560: gap of 100 bp
* 12561 13372: contig of 812 bp in length
* 13373 13472: gap of 100 bp
* 13473 14275: contig of 803 bp in length
* 14276 14375: gap of 100 bp
* 14376 15173: contig of 798 bp in length
* 15174 15273: gap of 100 bp
* 15274 16058: contig of 785 bp in length
* 16059 16158: gap of 100 bp
* 16159 16953: contig of 795 bp in length
* 16954 17053: gap of 100 bp
* 17054 17840: contig of 787 bp in length
* 17841 17940: gap of 100 bp
* 17941 18750: contig of 810 bp in length
* 18751 18850: gap of 100 bp
* 18851 19656: contig of 806 bp in length
* 19657 19756: gap of 100 bp
* 19757 20561: contig of 805 bp in length
* 20562 20661: gap of 100 bp
* 20662 21456: contig of 795 bp in length
* 21457 21556: gap of 100 bp
* 21557 22382: contig of 826 bp in length
* 22383 22482: gap of 100 bp
* 22483 23306: contig of 824 bp in length
* 23307 23406: gap of 100 bp
* 23407 24217: contig of 811 bp in length
* 24218 24317: gap of 100 bp
* 24318 25116: contig of 799 bp in length
* 25117 25216: gap of 100 bp
* 25217 25999: contig of 783 bp in length
* 26000 26099: gap of 100 bp
* 26100 26879: contig of 780 bp in length
* 26880 26979: gap of 100 bp
* 26980 27773: contig of 794 bp in length
* 27774 27873: gap of 100 bp
* 27874 28668: contig of 795 bp in length
* 28669 28768: gap of 100 bp
* 28769 29592: contig of 824 bp in length
* 29593 29692: gap of 100 bp
* 29693 30523: contig of 831 bp in length
* 30524 30623: gap of 100 bp
* 30624 31423: contig of 800 bp in length
* 31424 31523: gap of 100 bp
* 31524 32321: contig of 798 bp in length

* 32322 32421: gap of 100 bp
* 32422 33244: contig of 823 bp in length
* 33245 33344: gap of 100 bp
* 33345 34172: contig of 828 bp in length
* 34173 34272: gap of 100 bp
* 34273 35079: contig of 807 bp in length
* 35080 35179: gap of 100 bp
* 35180 35967: contig of 788 bp in length
* 35968 36067: gap of 100 bp
* 36068 36846: contig of 779 bp in length
* 36847 36946: gap of 100 bp
* 36947 37740: contig of 794 bp in length
* 37741 37840: gap of 100 bp
* 37841 38651: contig of 811 bp in length
* 38652 38751: gap of 100 bp
* 38752 39566: contig of 815 bp in length
* 39567 39666: gap of 100 bp
* 39667 40472: contig of 806 bp in length
* 40473 40572: gap of 100 bp
* 40573 41381: contig of 809 bp in length
* 41382 41481: gap of 100 bp
* 41482 42296: contig of 815 bp in length
* 42297 42396: gap of 100 bp
* 42397 43213: contig of 817 bp in length
* 43214 43313: gap of 100 bp
* 43314 44124: contig of 811 bp in length
* 44125 44224: gap of 100 bp
* 44225 45028: contig of 804 bp in length
* 45029 45128: gap of 100 bp
* 45129 45937: contig of 809 bp in length
* 45938 46037: gap of 100 bp
* 46038 46814: contig of 777 bp in length
* 46815 46914: gap of 100 bp
* 46915 47710: contig of 796 bp in length
* 47711 47810: gap of 100 bp
* 47811 48580: contig of 770 bp in length
* 48581 48680: gap of 100 bp
* 48681 49465: contig of 785 bp in length
* 49466 49565: gap of 100 bp
* 49566 50377: contig of 812 bp in length
* 50378 50477: gap of 100 bp
* 50478 51260: contig of 783 bp in length
* 51261 51360: gap of 100 bp
* 51361 52143: contig of 783 bp in length
* 52144 52243: gap of 100 bp
* 52244 53049: contig of 806 bp in length
* 53050 53149: gap of 100 bp
* 53150 53959: contig of 810 bp in length
* 53960 54059: gap of 100 bp
* 54060 54832: contig of 773 bp in length
* 54833 54932: gap of 100 bp
* 54933 55712: contig of 780 bp in length
* 55713 55812: gap of 100 bp
* 55813 56604: contig of 792 bp in length
* 56605 56704: gap of 100 bp
* 56705 57487: contig of 783 bp in length
* 57488 57587: gap of 100 bp
* 57588 58395: contig of 808 bp in length
* 58396 58495: gap of 100 bp
* 58496 59302: contig of 807 bp in length
* 59303 59402: gap of 100 bp
* 59403 60205: contig of 803 bp in length
* 60206 60305: gap of 100 bp
* 60306 61119: contig of 814 bp in length
* 61120 61219: gap of 100 bp
* 61220 62026: contig of 807 bp in length
* 62027 62126: gap of 100 bp
* 62127 62941: contig of 815 bp in length
* 62942 63041: gap of 100 bp
* 63042 63857: contig of 816 bp in length
* 63858 63957: gap of 100 bp
* 63958 64772: contig of 815 bp in length
* 64773 64872: gap of 100 bp

* 64873 65675: contig of 803 bp in length

Query Match 4.4%; Score 34; DB 2; Length 81953;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 88 GATCCAACTGGGAGTTGAAGTGTGAGTGAGACTG 121
|||||
Db 56024 GATCCAACTGGGAGTTGAAGTGTGAGTGAGACTG 56057

RESULT 15
AC002415 93419 bp DNA linear HTG 12-AUG-1997
LOCUS Homo sapiens chromosome X clone BWXD142, *** SEQUENCING IN PROGRESS
DEFINITION ***; 6 unordered pieces.
ACCESSION AC002415
VERSION AC002415.1 GI:2323256
KEYWORDS HTG; HTGS_PHASE1.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 93419)
AUTHORS Chen,E., Brownstein,B.H., States,D.J., Schlessinger,D. and Mazzarella,R.
TITLE Direct Submission
JOURNAL Unpublished (1997)
REFERENCE 2 (bases 1 to 93419)
AUTHORS Brownstein,B.H., States,D.J. and Mazzarella,R.
TITLE Direct Submission
JOURNAL Submitted (12-AUG-1997) Center for Genetics in Medicine, Box 8232, Washington University School of Medicine, 4566 Scott Avenue, St. Louis, MO 63110, USA
COMMENT Current status of this project is available at:
'http://genome.wustl.edu/cgm/seq_projects.html'
Submitted by:
Elison Chen,
Advanced Center for Genetic Technology,
Applied Biosystems Division of Perlin Elmer Corp.,
850 Lincoln Center Drive,
Foster City, CA 94404 USA
e-mail: ellison@genseq.apdbio.com

and

Buddy Brownstein,
Center for Genetics in Medicine,
Washington University School of Medicine, Box 8232
4566 Scott Avenue,
St. Louis, MO 63110, USA
e-mail: buddy@genetics.wustl.edu

and

David J. States,
Institute for Biomedical Computing
Washington University in St. Louis
700 South Euclid Ave.
St. Louis, MO 63108 USA
e-mail: states@ibc.wustl.edu.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 9560: contig of 9560 bp in length
* 9561 25127: contig of 15567 bp in length
* 25128 45675: contig of 20548 bp in length
* 45676 73399: contig of 27724 bp in length

* 73400 82075: contig of 8676 bp in length
* 82076 93419: contig of 11344 bp in length.
Location/Qualifiers
1. 93419
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/clone="BXND142"
BASE COUNT 25991 a 21189 c 20209 g 26030 t
ORIGIN

Query Match 4.4%; Score 34; DB 2; Length 93419;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 88 GATCCAACTGGGAGTTGAAGTGTGAGTGAGACTG 121
|||||
Db 55823 GATCCAACTGGGAGTTGAAGTGTGAGTGAGACTG 55856

RESULT 16
AL445227 114344 bp DNA linear HTG 10-JUL-2001
LOCUS Homo sapiens chromosome X clone RP13-97115, *** SEQUENCING IN
DEFINITION PROGRESS ***; 6 unordered pieces.
ACCESSION AL445227
VERSION AL445227.7 GI:11611395
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 114344)
AUTHORS McIay,K.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Dec 9, 2000 this sequence version replaced gi:11493333.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: b97115
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 112482 bases at least Q40
Consensus quality: 113245 bases at least Q30
Consensus quality: 113597 bases at least Q20
Insert size: 113844; sum-of-contigs
Quality coverage: 5.91x in Q20 bases; agarose-fp
coverage: 5.72x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 13809: contig of 13809 bp in length
* 13810 13909: gap of 100 bp
* 13910 51507: contig of 37598 bp in length
* 51508 51607: gap of 100 bp
* 51608 65679: contig of 14072 bp in length
* 65680 65779: gap of 100 bp
* 65780 69064: contig of 3285 bp in length

repeat_region /note="MIR repeat: matches 98. .247 of consensus"
14083. .14381
repeat_region /note="AluX repeat: matches 1. .293 of consensus"
14906. .15126
repeat_region /note="L2 repeat: matches 2364. .2569 of consensus"
15142. .15370
repeat_region /note="MER4B repeat: matches 1. .215 of consensus"
15366. .16028
repeat_region /note="MER4D repeat: matches 359. .1017 of consensus"
16097. .16211
repeat_region /note="L2 repeat: matches 2631. .2749 of consensus"
16199. .16240
repeat_region /note="L2 repeat: matches 2560. .2599 of consensus"
18264. .18349
repeat_region /note="43 copies 2 mer ga 62% conserved"
19489. .19797
repeat_region /note="AluJb repeat: matches 1. .306 of consensus"
19826. .20044
repeat_region /note="L1M3 repeat: matches 5220. .5419 of consensus"
20045. .20175
repeat_region /note="FLAM_A repeat: matches 1. .131 of consensus"
20176. .20249
repeat_region /note="L1M3 repeat: matches 5419. .5496 of consensus"
20270. .20929
repeat_region /note="L1MA5A repeat: matches 5447. .6104 of consensus"
20930. .21226
repeat_region /note="AluX repeat: matches 1. .297 of consensus"
21227. .21417
repeat_region /note="L1MA5A repeat: matches 6104. .6294 of consensus"
21418. .21461
repeat_region /note="L1M4 repeat: matches 4295. .4338 of consensus"
21500. .23928
repeat_region /note="L1PA5 repeat: matches 3696. .6143 of consensus"
24145. .24341
repeat_region /note="L1MEC repeat: matches 1820. .2012 of consensus"
24385. .24681
repeat_region /note="AluX repeat: matches 1. .296 of consensus"
24718. .28764
repeat_region /note="L1PA4 repeat: matches 2108. .6144 of consensus"
28765. .28886
repeat_region /note="FLAM_C repeat: matches 3. .124 of consensus"
28894. .28928
repeat_region /note="L1PA4 repeat: matches 2064. .2097 of consensus"
28929. .29221
repeat_region /note="AluSc repeat: matches 1. .296 of consensus"
29222. .29377
repeat_region /note="L1MEC repeat: matches 2097. .2258 of consensus"
29428. .29551
repeat_region /note="L2 repeat: matches 2621. .2749 of consensus"
29553. .29644
repeat_region /note="AluJb repeat: matches 205. .296 of consensus"
29648. .30097
repeat_region /note="MER87 repeat: matches 1. .469 of consensus"
30104. .30153
repeat_region /note="25 copies 2 mer ca 76% conserved"
31811. .31976
repeat_region /note="MIR repeat: matches 46. .204 of consensus"
32005. .32260
repeat_region /note="128 copies 2 mer ta 80% conserved"
32903. .33194
repeat_region /note="AluJo repeat: matches 1. .293 of consensus"
33677. .34221
repeat_region /note="MER68A repeat: matches 1. .568 of consensus"
34372. .35093
repeat_region /note="TIGGER1 repeat: matches 1650. .2308 of consensus"
35094. .35333
repeat_region /note="L1PB1 repeat: matches 5909. .6155 of consensus"
35334. .35437
repeat_region /note="TIGGER1 repeat: matches 2308. .2418 of consensus"
35477. .35878
repeat_region /note="MLT1-INTERNAL repeat: matches 923. .1579 of consensus"
36082. .36740

repeat_region /note="MLT1-INTERNAL repeat: matches 201. .885 of consensus"
36783. .37121
repeat_region /note="L2 repeat: matches 824. .1219 of consensus"
37158. .37328
repeat_region /note="L1M4 repeat: matches 4640. .4816 of consensus"
39953. .39978
repeat_region /note="13 copies 2 mer ca 100% conserved"
40525. .40566
repeat_region /note="MIR repeat: matches 110. .151 of consensus"
44063. .44377
repeat_region /note="AluJo repeat: matches 1. .299 of consensus"
44451. .45182
repeat_region /note="MER21B repeat: matches 3. .790 of consensus"
45243. .45477
repeat_region /note="L1M4 repeat: matches 4078. .4321 of consensus"
45515. .45890
repeat_region /note="L1PA7 repeat: matches 5769. .6143 of consensus"
45900. .47783
repeat_region /note="L1PA8 repeat: matches 4284. .6161 of consensus"
47776. .48115
repeat_region /note="L1M4 repeat: matches 3680. .4029 of consensus"
48116. .48477
repeat_region /note="MER47A repeat: matches 1. .365 of consensus"
48478. .48554

Query Match 4.4% Score 34; DB 9; Length 149749;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GATCCAACTGGAGTTGAAGTGTGAGTGAGAGTG 121
|||||
Db 95143 GATCCAACTGGAGTTGAAGTGTGAGTGAGAGTG 95110

RESULT 18
AC002415/c
LOCUS
DEFINITION Homo sapiens chromosome X clone BWXD142, *** SEQUENCING IN PROGRESS
AC002415 93419 bp DNA linear HTG 12-AUG-1997
***, 6 unordered pieces.
AC002415 AC002415.J GI:2323256
VERSION HTG; HTGS_PHASE1.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM
REFERENCE 1 (bases 1 to 93419)
AUTHORS Chen, E., Brownstein, B.H., States, D.J., Schlessinger, D. and Mazzarella, R.
TITLE Direct Submission
JOURNAL Unpublished (1997)
REFERENCE 2 (bases 1 to 93419)
AUTHORS Brownstein, B.H., States, D.J. and Mazzarella, R.
TITLE Direct Submission
JOURNAL Submitted (12-AUG-1997) Center for Genetics in Medicine, Box 8232, Washington University School of Medicine, 4566 Scott Avenue, St. Louis, MO 63110, USA
COMMENT Current status of this project is available at: 'http://genome.wustl.edu/cgm/seq_projects.html', Submitted by: Ellson Chen, Advanced Center for Genetic Technology, Applied Biosystems Division of Perlin Elmer Corp., 850 Lincoln Center Drive, Foster City, CA 94404 USA e-mail: ellson@genseq.apldbio.com and Buddy Brownstein, Center for Genetics in Medicine, Washington University School of Medicine, Box 8232

4566 Scott Avenue,
St. Louis, MO 63110, USA
e-mail: buddygenetics.wustl.edu
and

David J. States,
Institute for Biomedical Computing
Washington University in St. Louis
700 South Euclid Ave.
St. Louis, MO 63108 USA

e-mail: states@bc.wustl.edu.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

* 1 9560: contig of 9560 bp in length
* 9561 25127: contig of 15567 bp in length
* 25128 45675: contig of 20548 bp in length
* 45676 73399: contig of 27724 bp in length
* 73400 82075: contig of 8676 bp in length
* 82076 93419: contig of 11344 bp in length.

FEATURES
source
1. .93419 Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/clone="BXND142"

BASE COUNT 25991 a 21189 c 20209 g 26030 t

Query Match 3.7%; Score 29; DB 2; Length 93419;
Best Local Similarity 100.0%; Pred. No. 0.0036;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 ATCCAACTGGGAGTTGAAGTGTGAGTGAG 117
|||||
Db 1844 ATCCAACTGGGAGTTGAAGTGTGAGTGAG 1816

RESULT 19
AL445227/c 114344 bp DNA linear HTG 10-JUL-2001
LOCUS Homo sapiens chromosome X clone RP13-97I15, *** SEQUENCING IN
DEFINITION PROGRESS ***, 6 unordered pieces.

AL445227
ACCESSION AL445227
VERSION AL445227.7 GI:11611395
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 114344)

McLay, K.
Direct Submission
Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Dec 9, 2000 this sequence version replaced gi:11493333.

----- Genome Center
Center: Sanger Centre
Center code: SC

Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

----- Project Information
Center project name: BB97I15
----- Summary Statistics

Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 112482 bases at least Q40
Consensus quality: 113245 bases at least Q30
Consensus quality: 113597 bases at least Q20
Insert size: 113844; sum-of-contigs
Insert size: 122396; 9.8% error; agarose-fp
Quality coverage: 5.91x in Q20 bases; sum-of-contigs Quality
coverage: 5.72x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 13809: contig of 13809 bp in length
* 13810 13909: gap of 100 bp
* 13910 51507: contig of 37598 bp in length
* 51508 51607: gap of 100 bp
* 51608 65679: contig of 14072 bp in length
* 65680 65779: gap of 100 bp
* 65780 69064: contig of 3285 bp in length
* 69065 69164: gap of 100 bp
* 69165 101518: contig of 32354 bp in length
* 101519 101618: gap of 100 bp
* 101619 114344: contig of 12726 bp in length.

FEATURES
source
1. .114344 Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/clone="RP13-97I15"
/clone_lib="RPC1-13.1"

misc_feature
1. .113809
/note="assembly_fragment:00241"
fragment_chain:1
clone_end:SP6
vector_side:left
13910..51507

misc_feature
/note="assembly_fragment:00680"
fragment_chain:1
51608..65679

misc_feature
/note="assembly_fragment:00207"
fragment_chain:1
65780..69064

misc_feature
/note="assembly_fragment:00300"
69165..101518

misc_feature
/note="assembly_fragment:01024"
101619..114344
/note="assembly_fragment:01825"
clone_end:T7
vector_side:right

BASE COUNT 32658 a 24338 c 24232 g 32616 t 500 others

ORIGIN

Query Match 3.7%; Score 29; DB 2; Length 114344;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 ATCCAACTGGGAGTTGAAGTGTGAGTGAG 117
|||||
Db 54734 ATCCAACTGGGAGTTGAAGTGTGAGTGAG 54706

RESULT 20

AL159987/c

LOCUS

DEFINITION

AL159987

Human DNA sequence from clone Rp11-167P23 on chromosome X, complete

sequence.

AL159987

ACCESSION

VERSION

AL159987.19 GI:13559998

KEYWORDS
SOURCE

HTG.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 119024)
Howden, P.
Direct Submission
Submitted (05-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

On Apr 6, 2001 this sequence version replaced gi:13274303.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em.: EMBL; Sw.:
SWISSPROT; Tr.: TREMBL; Wp.: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C-elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome X, constructed by the Sanger Centre Chromosome X Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/chrX
RP11-167P23 is from the library RP11-11.1 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
IMPORTANT: This sequence is not the entire insert of clone
RP11-167P23 It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true left end of clone RP11-465E19 is at 118925 in this
sequence. The true right end of clone RP11-47K8 is at 100 in this
sequence.

FEATURES

Source

Location/Qualifiers
1..119024
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/clone="RP11-167P23"
/clone_lib="RPC11.1"
1..1665

repeat_region

/note="HERVL repeat: matches 1583..3249 of consensus"
1704..1743
/note="20 copies 2 mer ac 97% conserved"
1744..2134
/note="MLT2B repeat: matches 1..392 of consensus"
2381..2705
/note="AluJo repeat: matches 1..312 of consensus"
3061..3437
/note="RHEIC repeat: matches 1..371 of consensus"
3445..3470
/note="13 copies 2 mer tg 92% conserved"
4743..4916
/note="MIR repeat: matches 7..195 of consensus"
5565..5682
/note="MIR repeat: matches 101..213 of consensus"
9092..9323
/note="match: GSS: Em:AQ980738"
9325..9686
/note="match: GSS: Em:AQ980738"
complement(9528..9691)
/note="match: STS: Em:G06170"

repeat_region

misc_feature

misc_feature

misc_feature

repeat_region

repeat_region

misc_feature

misc_feature

misc_feature

misc_feature

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

```
misc_feature      /note="match: GSS: Em:B47902"
32364..32727
/note="match: GSS: Em:AQ107228"
32412..32777
/note="L1PA8 repeat: matches 5793..6159 of consensus"
32835..33112
/note="L2 repeat: matches 2255..2522 of consensus"
33291..33584
/note="MER39 repeat: matches 98..395 of consensus"
34330..34453
/note="FLAM_C repeat: matches 1..124 of consensus"
34597..34970
/note="LTR3 repeat: matches 1..413 of consensus"
complement(join(35694..36254,36581..36721))
/note="match: GSS: Em:AQ389427"
36267..36573
/note="ALUSC repeat: matches 1..309 of consensus"
complement(36574..36730)
/note="match: GSS: Em:AQ181456"
38336..38377
/note="MIR repeat: matches 110..151 of consensus"
complement(38817..39110)
/note="match: STS: Em:HS220XG1"
41596..41706
/note="L1 repeat: matches 4663..4779 of consensus"
41765..41821
/note="L2 repeat: matches 1163..1219 of consensus"
41849..42117
/note="L2 repeat: matches 811..1102 of consensus"
42146..42812
/note="MLT1-INTERNAL repeat: matches 201..885 of
consensus"
43013..43410
/note="MLT1-INTERNAL repeat: matches 920..1322 of
consensus"
43499..43559
/note="TIGER1 repeat: matches 2288..2359 of consensus"
43560..43799
/note="L1PB1 repeat: matches 5909..6155 of consensus"
43800..44089
/note="TIGER1 repeat: matches 2078..2288 of consensus"
44132..44550
/note="TIGER1 repeat: matches 1650..2094 of consensus"
44701..45249
/note="MER68A repeat: matches 1..568 of consensus"
45741..46034
/note="AluJo repeat: matches 1..293 of consensus"
46311..46342
/note="16 copies 2 mer ct 90% conserved"

Query Match      3.7%; Score 29; DB 9; Length 119024;
Best Local Similarity 100.0%; Pred. No. 0.0033;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 ATCCAACTGGGAGTTCGAAGTGTGAGTGAG 117
|||||
Db 18649 ATCCAACTGGGAGTTCGAAGTGTGAGTGAG 18621

RESULT 21
AL445236
LOCUS      AL445236      149749 bp      DNA      linear      PRI 09-MAR-2001
DEFINITION Human DNA sequence from clone RP13-77011 on chromosome
Xp11.21-11.3, complete sequence.
ACCESSION  AL445236
VERSION    AL445236.22  GI:13184376
KEYWORDS   HTG.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 149749)
AUTHORS   Howden,P.
```

```

repeat_region 9015..9136
repeat_region /note="61 copies 2 mer ct 69% conserved"
repeat_region 9677..9718
repeat_region /note="21 copies 2 mer tt 85% conserved"
repeat_region 9745..10036
repeat_region /note="MER92C repeat: matches 272..552 of consensus"
repeat_region 10551..10751
repeat_region /note="MER92B repeat: matches 2..212 of consensus"
repeat_region 10857..10969
repeat_region /note="MIR repeat: matches 40..154 of consensus"
repeat_region 11521..11613
repeat_region /note="MIR repeat: matches 63..156 of consensus"
repeat_region 11629..12187
repeat_region /note="L2 repeat: matches 1779..2341 of consensus"
repeat_region 12438..12548
repeat_region /note="L1MA9 repeat: matches 6163..6270 of consensus"
repeat_region 12975..13093
repeat_region /note="MIR repeat: matches 98..247 of consensus"
repeat_region 14083..14381
repeat_region /note="AluX repeat: matches 1..293 of consensus"
repeat_region 14906..15126
repeat_region /note="L2 repeat: matches 2364..2569 of consensus"
repeat_region 15142..15370
repeat_region /note="MER4B repeat: matches 1..215 of consensus"
repeat_region 15366..16028
repeat_region /note="MER4D repeat: matches 359..1017 of consensus"
repeat_region 16097..16211
repeat_region /note="L2 repeat: matches 2631..2749 of consensus"
repeat_region 16199..16240
repeat_region /note="L2 repeat: matches 2560..2599 of consensus"
repeat_region 18264..18349
repeat_region /note="43 copies 2 mer ga 62% conserved"
repeat_region 19489..19797
repeat_region /note="AluJ repeat: matches 1..306 of consensus"
repeat_region 19826..20044
repeat_region /note="L1M3 repeat: matches 5220..5419 of consensus"
repeat_region 20045..20175
repeat_region /note="FLAM.A repeat: matches 1..131 of consensus"
repeat_region 20176..20249
repeat_region /note="L1M3 repeat: matches 5419..5496 of consensus"
repeat_region 20270..20929
repeat_region /note="L1MA5A repeat: matches 5447..6104 of consensus"
repeat_region 20930..21226
repeat_region /note="AluX repeat: matches 1..297 of consensus"
repeat_region 21227..21417
repeat_region /note="L1MA5A repeat: matches 6104..6294 of consensus"
repeat_region 21418..21461
repeat_region /note="L1M4 repeat: matches 4295..4338 of consensus"
repeat_region 21500..23928
repeat_region /note="L1PA5 repeat: matches 3696..6143 of consensus"
repeat_region 24145..24341
repeat_region /note="L1MEC repeat: matches 1820..2012 of consensus"
repeat_region 24385..24681
repeat_region /note="AluX repeat: matches 1..296 of consensus"
repeat_region 24718..28764
repeat_region /note="L1PA4 repeat: matches 2108..6144 of consensus"
repeat_region 28765..28886
repeat_region /note="FLAM.C repeat: matches 3..124 of consensus"
repeat_region 28894..28928
repeat_region /note="L1PA4 repeat: matches 2064..2097 of consensus"
repeat_region 28929..29221
repeat_region /note="AluSc repeat: matches 1..296 of consensus"
repeat_region 29222..29377
repeat_region /note="L1MEC repeat: matches 2097..2258 of consensus"
repeat_region 29428..29551
repeat_region /note="L2 repeat: matches 2621..2749 of consensus"
repeat_region 29553..29644
repeat_region /note="AluJ repeat: matches 205..296 of consensus"
repeat_region 29648..30097
repeat_region /note="MER87 repeat: matches 1..469 of consensus"
repeat_region 30104..30153
repeat_region /note="25 copies 2 mer ca 76% conserved"
repeat_region 31811..31976

```

```

repeat_region /note="MIR repeat: matches 46..204 of consensus"
repeat_region 32005..32260
repeat_region /note="128 copies 2 mer ta 80% conserved"
repeat_region 32903..33194
repeat_region /note="AluJ repeat: matches 1..293 of consensus"
repeat_region 33677..34221
repeat_region /note="MER68A repeat: matches 1..568 of consensus"
repeat_region 34372..35093
repeat_region /note="TIGGER1 repeat: matches 1650..2308 of consensus"
repeat_region 35094..35333
repeat_region /note="L1PB1 repeat: matches 5909..6155 of consensus"
repeat_region 35334..35437
repeat_region /note="TIGGER1 repeat: matches 2308..2418 of consensus"
repeat_region 35477..35878
repeat_region /note="MULTI-INTERNAL repeat: matches 923..1579 of consensus"
repeat_region 36082..36740
repeat_region /note="MULTI-INTERNAL repeat: matches 201..885 of consensus"
repeat_region 36783..37121
repeat_region /note="L2 repeat: matches 824..1219 of consensus"
repeat_region 37158..37328
repeat_region /note="L1M4 repeat: matches 4640..4816 of consensus"
repeat_region 39953..39978
repeat_region /note="13 copies 2 mer ca 100% conserved"
repeat_region 40525..40566
repeat_region /note="MIR repeat: matches 110..151 of consensus"
repeat_region 44063..44377
repeat_region /note="AluJ repeat: matches 1..299 of consensus"
repeat_region 44451..45182
repeat_region /note="MER21B repeat: matches 3..790 of consensus"
repeat_region 45243..45477
repeat_region /note="L1M4 repeat: matches 4078..4321 of consensus"
repeat_region 45515..45890
repeat_region /note="L1PA7 repeat: matches 5769..6143 of consensus"
repeat_region 45900..47783
repeat_region /note="L1PA8 repeat: matches 4284..6161 of consensus"
repeat_region 47776..48115
repeat_region /note="L1M4 repeat: matches 3680..4029 of consensus"
repeat_region 48116..48477
repeat_region /note="MER47A repeat: matches 1..365 of consensus"
repeat_region 48478..48554

```

Query Match 3.7%; Score 29; DB 9; Length 149749;
 Best Local Similarity 100.0%; Pred. No. 0.0031;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 ATCCAACTGGGAGTTGAAGTGTGAGTGAG 117
 |||||

Db 76369 ATCCAACTGGGAGTTGAAGTGTGAGTGAG 76397

RESULT 22
 HSA318881 HSA318881 493 bp mRNA linear PRI 14-JAN-2002
 LOCUS Homo sapiens mRNA for XAGE-3 protein.
 DEFINITION
 ACCESSION AJ318881
 VERSION AJ318881.1 GI:18157211
 KEYWORDS XAGE-3 gene.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (slices)
 AUTHORS Zendman,A.J.W., van Kraats,A.A., Weidie,U.H., Ruiters,D.R. and Van
 Muijen,G.N.P.
 Expression profile of members of the XAGE cancer/testis antigen
 family
 TITLE
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 493)
 AUTHORS Zendman,A.J.W.
 TITLE Direct Submission
 JOURNAL Submitted (14-AUG-2001) Zendman A.J.W., Department of Pathology,

University Medical Center St Radboud, Geert Grooteplein Zuid 24, PO
Box 9101 6500 HB Nijmegen, NETHERLANDS

FEATURES

source
Location/Qualifiers
1..493
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp11.21-Xp11.22"
/tissue.type="Ewing sarcoma"
61..396
/gene="XAGE-3"
61..396
/gene="XAGE-3"
/codon_start=1
/product="XAGE-3 protein"
/protein_id="CAC83008.1"
/db_xref="GI:18157212"
/translation="MIWGRSTYRPRRSVPPPELIGPMLPGDEEPQOEPPPTESR
DPAPQEREDQGAETQVPDLEADLQELSQSKTGCGNGPDDQGKILPKSEQFKMP
EGDQRPQV"

gene
CDS
167 a 98 c 125 g 103 t

Query Match 3.5%; Score 27; DB 9; Length 493;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 679 GTTTAAATGAAGACAGCTGAACAAC 705
|||||
Db 391 GTTTAAATGAAGACAGCTGAACAAC 417
|||||

RESULT 23
HSA318893
LOCUS
DEFINITION Homo sapiens XAGE-3 gene, exons 1-5.
ACCESSION AJ318893
VERSION AJ318893.1 GI:18181827
KEYWORDS XAGE-3 gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
Zendman,A.J.W., van Kraats,A.A., Weidie,U.H., Ruiter,D.J. and Van
Muijen,G.N.P.
Expression profile and alignment of the XAGE family of
cancer/testis associated genes
Unpublished
2 (bases 1 to 7200)
Zendman,A.J.W.
Direct Submission
Submitted (15-AUG-2001) Zendman A.J.W., Department of Pathology,
University Medical Center, Geert Grooteplein Zuid 24, P.O. Box 9101
6500 HB Nijmegen, NETHERLANDS
Location/Qualifiers
1..7200
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp11.21-Xp11.22"
join(1282..1333,1442..1530,2045..2150,3683..3808,
5943..6052)
/gene="XAGE-3"
1282..6052
/gene="XAGE-3"
1282..1333
/gene="XAGE-3"
/number=1
1334..1441
/gene="XAGE-3"
/number=1
1442..1530

BASE COUNT 167 a 98 c 125 g 103 t
EGDQRPQV"

Query Match 3.5%; Score 27; DB 9; Length 493;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 679 GTTTAAATGAAGACAGCTGAACAAC 705
|||||
Db 391 GTTTAAATGAAGACAGCTGAACAAC 417
|||||

RESULT 24
AL590240
LOCUS
DEFINITION Human DNA sequence from clone RP11-26613 on chromosome X, complete
sequence.
ACCESSION AL590240 AC025380
VERSION AL590240.5 GI:13872461
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 130005)
Clark,S.
Direct Submission
Submitted (25-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On May 15, 2001 this sequence version replaced gi:8705130
gi:13785125.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30);
an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the

BASE COUNT 2131 a 1327 c 1573 g 2169 t
ORIGIN

Query Match 3.5%; Score 27; DB 9; Length 7200;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 679 GTTTAAATGAAGACAGCTGAACAAC 705
|||||
Db 5960 GTTTAAATGAAGACAGCTGAACAAC 5986
|||||

RESULT 24
AL590240
LOCUS
DEFINITION Human DNA sequence from clone RP11-26613 on chromosome X, complete
sequence.
ACCESSION AL590240 AC025380
VERSION AL590240.5 GI:13872461
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 130005)
Clark,S.
Direct Submission
Submitted (25-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On May 15, 2001 this sequence version replaced gi:8705130
gi:13785125.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30);
an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the

CDs
/gene="XAGE-3"
join(1450..1530,2045..2150,3683..3808,5943..5965)
/gene="XAGE-3"
/codon_start=1
/product="XAGE-3 protein"
/protein_id="CAC83014.1"
/db_xref="GI:18181828"
/translation="MIWGRSTYRPRRSVPPPELIGPMLPGDEEPQOEPPPTESR
DPAPQEREDQGAETQVPDLEADLQELSQSKTGCGNGPDDQGKILPKSEQFKMP
EGDQRPQV"
1531..2044
/gene="XAGE-3"
/number=2
2045..2150
/gene="XAGE-3"
/number=3
2151..3682
/gene="XAGE-3"
/number=3
3683..3808
/gene="XAGE-3"
/number=4
3809..5942
/gene="XAGE-3"
/number=4
5943..6052
/gene="XAGE-3"
/number=5

intron
exon
intron
exon
intron
exon
BASE COUNT 2131 a 1327 c 1573 g 2169 t
ORIGIN

Query Match 3.5%; Score 27; DB 9; Length 7200;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 679 GTTTAAATGAAGACAGCTGAACAAC 705
|||||
Db 5960 GTTTAAATGAAGACAGCTGAACAAC 5986
|||||

RESULT 24
AL590240
LOCUS
DEFINITION Human DNA sequence from clone RP11-26613 on chromosome X, complete
sequence.
ACCESSION AL590240 AC025380
VERSION AL590240.5 GI:13872461
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 130005)
Clark,S.
Direct Submission
Submitted (25-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On May 15, 2001 this sequence version replaced gi:8705130
gi:13785125.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30);
an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the

BASE COUNT 2131 a 1327 c 1573 g 2169 t
ORIGIN

Query Match 3.5%; Score 27; DB 9; Length 7200;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 679 GTTTAAATGAAGACAGCTGAACAAC 705
|||||
Db 5960 GTTTAAATGAAGACAGCTGAACAAC 5986
|||||

RESULT 24
AL590240
LOCUS
DEFINITION Human DNA sequence from clone RP11-26613 on chromosome X, complete
sequence.
ACCESSION AL590240 AC025380
VERSION AL590240.5 GI:13872461
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 130005)
Clark,S.
Direct Submission
Submitted (25-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On May 15, 2001 this sequence version replaced gi:8705130
gi:13785125.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30);
an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/ChrX> RP11-26613 is from the library RP11-11.1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBACe3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-26613 It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.

The true left end of clone RP11-382F24 is at 129906 in this sequence. The true right end of clone RPI-296K21 is at 100 in this sequence.

FEATURES

source	Location/Qualifiers
repeat_region	1..130005 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="X" /clone="RP11-26613" /clone_lib="RP11-11.1" 1..3675
repeat_region	/note="L1P repeat: matches 4..3892 of consensus"
repeat_region	4031..4472 /note="L1PA10 repeat: matches 5718..6165 of consensus"
repeat_region	4477..4658 /note="L1ME3A repeat: matches 5967..6157 of consensus"
repeat_region	4869..5170 /note="AluSc repeat: matches 1..306 of consensus"
repeat_region	6010..6307 /note="AluX repeat: matches 1..299 of consensus"
repeat_region	9580..9670 /note="MIR repeat: matches 127..214 of consensus"
repeat_region	10019..10387 /note="L1T1B repeat: matches 41..390 of consensus"
repeat_region	10424..11726 /note="L1W7 repeat: matches 4980..6284 of consensus"
repeat_region	12183..12372 /note="L1PA15 repeat: matches 5963..6153 of consensus"
repeat_region	12387..17012 /note="L1MA2 repeat: matches 1626..6308 of consensus"
repeat_region	17013..17325 /note="AluSq repeat: matches 1..313 of consensus"
repeat_region	17326..17971 /note="L1MA2 repeat: matches 979..1626 of consensus"
repeat_region	18003..24449 /note="L1PA7 repeat: matches 6..6143 of consensus"
repeat_region	24453..24926 /note="L1P repeat: matches 4190..4658 of consensus"
repeat_region	25188..25377 /note="L1M4 repeat: matches 4619..4820 of consensus"
repeat_region	27242..27405 /note="82 copies 2 mer ta 76% conserved"
repeat_region	27731..27937 /note="L1M4 repeat: matches 2..231 of consensus"
repeat_region	28628..28710 /note="L1T9/FLAM repeat: matches 2..84 of consensus"
repeat_region	28760..28926 /note="L1MA8 repeat: matches 6114..6290 of consensus"
repeat_region	30040..30466 /note="L1T9 repeat: matches 198..625 of consensus"
repeat_region	30467..31208 /note="L1 repeat: matches 3291..4031 of consensus"
repeat_region	31209..31480 /note="AluX repeat: matches 38..311 of consensus"
repeat_region	31481..32418 /note="L1 repeat: matches 2350..3291 of consensus"
repeat_region	32439..33203 /note="L1PA13 repeat: matches 5395..6144 of consensus"
repeat_region	33211..33753 /note="L1P repeat: matches 1840..2380 of consensus"
repeat_region	33748..35989 /note="L1P repeat: matches 2006..4258 of consensus"
repeat_region	35992..36688 /note="L1PA3 repeat: matches 5449..6146 of consensus"
repeat_region	36889..36728 /note="20 copies 2 mer aa 80% conserved"
repeat_region	36730..37022 /note="L1P repeat: matches 1564..1858 of consensus"
repeat_region	37018..37317 /note="L1P4 repeat: matches 5312..5604 of consensus"
repeat_region	37327..37629 /note="AluX repeat: matches 1..303 of consensus"
repeat_region	37630..37763 /note="AluSg1 repeat: matches 2..133 of consensus"
repeat_region	37782..38928 /note="L1P4 repeat: matches 4190..5309 of consensus"
repeat_region	39179..39381 /note="L1M4 repeat: matches 4604..4820 of consensus"
misc_feature	complement(40796..41371) /note="match: STS: Em:L30416"
repeat_region	41220..41301 /note="41 copies 2 mer tc 76% conserved"
misc_feature	41260..41339 /note="match: STS: Em:G09960 Em:G34497"
repeat_region	42561..42643 /note="AluJ/FLAM repeat: matches 2..84 of consensus"
repeat_region	42693..42859 /note="L1MA8 repeat: matches 6114..6290 of consensus"
repeat_region	43973..44399 /note="L1T9 repeat: matches 198..625 of consensus"
repeat_region	44400..44975 /note="L1P4 repeat: matches 3453..4031 of consensus"
repeat_region	44980..46869 /note="L1P3 repeat: matches 246..2016 of consensus"
repeat_region	47559..47715 /note="MIR repeat: matches 59..216 of consensus"
repeat_region	48511..48672 /note="L2 repeat: matches 2598..2750 of consensus"
repeat_region	48814..48941 /note="L1MD repeat: matches 177..303 of consensus"
repeat_region	49219..49248 /note="15 copies 2 mer tg 86% conserved"
repeat_region	49262..49293 /note="16 copies 2 mer ta 87% conserved"
repeat_region	49387..49410 /note="12 copies 2 mer ac 100% conserved"
repeat_region	49711..50182 /note="L2 repeat: matches 2252..2706 of consensus"
repeat_region	50693..50992 /note="AluY repeat: matches 10..299 of consensus"
repeat_region	51042..51248 /note="MIR repeat: matches 44..258 of consensus"
repeat_region	51506..51944 /note="L1MD2 repeat: matches 5686..6131 of consensus"
repeat_region	51943..52075 /note="L1MD3 repeat: matches 7590..7739 of consensus"
repeat_region	53350..53652 /note="AluSc repeat: matches 1..308 of consensus"
repeat_region	54185..54326 /note="MIR repeat: matches 117..261 of consensus"
repeat_region	54859..55936 /note="L1M1 repeat: matches 1825..2881 of consensus"
repeat_region	55971..56980 /note="L1PA5 repeat: matches 5130..6138 of consensus"
repeat_region	56977..57337 /note="L1P repeat: matches 4625..4985 of consensus"
repeat_region	57369..57577 /note="L1 repeat: matches 2958..3167 of consensus"
repeat_region	57547..57704

```
/note="L1 repeat: matches 4465. .4628 of consensus"
57701. .58142
/note="L1MA4 repeat: matches 5662. .6185 of consensus"
58673. .58774
/note="L1MA8 repeat: matches 5573. .5661 of consensus"
58964. .59010
/note="L2 repeat: matches 1865. .1911 of consensus"
59415. .59702
/note="L1MA8 repeat: matches 5778. .6073 of consensus"
59704. .59857
/note="MER41B repeat: matches 482. .635 of consensus"
59860. .60918
/note="MER41-internal repeat: matches 3. .1717 of
consensus"
60942. .61207
/note="MER41-internal repeat: matches 985. .1277 of
consensus"
61204. .61802
/note="MER41-internal repeat: matches 3340. .3944 of
consensus"
61819. .62285
/note="MER41A repeat: matches 1. .467 of consensus"
62289. .62468
/note="L1MA8 repeat: matches 6059. .6243 of consensus"

Query Match      3.2%; Score 25; DB 9; Length 130005;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 469 CTGAAAGTCGGGATCCTACACCTGG 493
|||||
Db 124239 CTGAAAGTCGGGATCCTACACCTGG 124263

RESULT 25
HSA318895      HSA318895      209 bp mRNA linear PRI 15-JAN-2002
LOCUS          Homo sapiens partial mRNA for XAGE-4 protein.
DEFINITION
ACCESSION      AJ318895
VERSION        AJ318895.1 GI:18181831
KEYWORDS       XAGE-4 gene.
SOURCE         human.
ORGANISM       Homo sapiens
REFERENCE      1 (sites)
AUTHORS        Zendenman,A.J.W., van Kraats,A.A., Weidie,U.H., Ruiter,D.J. and Van
Muijen,G.N.P.
TITLE          Expression profile and alignment of the XAGE family of
cancer/testis associated genes
JOURNAL        Unpublished
REFERENCE      2 (bases 1 to 209)
AUTHORS        Zendenman,A.J.W.
TITLE          Direct Submission
JOURNAL        Submitted (15-AUG-2001) Zendenman A.J.W., Department of Pathology,
University Medical Center, Geert Grooteplein Zuid 24, P.O. Box 9101
6500 HB Nijmegen, NETHERLANDS
FEATURES       Location/Qualifiers
source         1. .209
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
gene           1. .209
               /gene="XAGE-4"
CDS            <1. .>209
               /gene="XAGE-4"
               /codon_start=3
               /product="XAGE-4 protein"
               /protein_id="CAC83092.1"
               /db_xref="GI:18181832"
               /translation="PPELIGPMLEPSDEEPQOEPPTESRDTTPVPDLTDLQELSQ
SKYDCEKRGDPDDKGKIQPNQSNLK"

BASE COUNT     64 a 51 c 55 g 39 t
ORIGIN
```

```
Query Match      3.1%; Score 24; DB 9; Length 209;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 469 CTGAAAGTCGGGATCCTACACCTG 492
|||||
Db 73 CTGAAAGTCGGGATCCTACACCTG 96

RESULT 26
AX246260      AX246260      391 bp DNA linear PAT 28-SEP-2001
LOCUS          Sequence 1190 from Patent WO0166753.
DEFINITION
ACCESSION      AX246260
VERSION        AX246260.1 GI:15860934
KEYWORDS
SOURCE         human.
ORGANISM       Homo sapiens
REFERENCE      1 (bases 1 to 391)
AUTHORS        Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D.,
Sidduth-Klinger,J., Reinhard,C., Randazzo,F., Kennedy,G.C., Pot,D.,
Kassam,A., Lamson,G., Drmanac,R., Crkvenjakov,R., Dickson,M.,
Drmanac,S., Labat,I., Leshkowitz,D., Kita,D., Garcia,V. and
Stache-Crain,B.
TITLE          Human genes and gene expression products
JOURNAL        Patent: WO 0166753-A 1190 13-SEP-2001;
Chiron Corporation (US) ; Hyseq Inc. (US)
FEATURES       Location/Qualifiers
source         1. .391
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
BASE COUNT     82 a 109 c 117 g 83 t
ORIGIN

Query Match      3.1%; Score 24; DB 6; Length 391;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 AGTGAGAGGCCCTCGAAGTCGTC 192
|||||
Db 169 AGTGAGAGGCCCTCGAAGTCGTC 192

RESULT 27
AL158819/c     AL158819      174913 bp DNA linear PRI 11-JAN-2001
LOCUS          Human DNA sequence from clone RP11-382F24 on chromosome X, complete
DEFINITION      sequence.
ACCESSION      AL158819
VERSION        AL158819.14 GI:12214278
KEYWORDS       HTG.
SOURCE         human.
ORGANISM       Homo sapiens
REFERENCE      1 (bases 1 to 174913)
AUTHORS        Direct Submission
JOURNAL        Submitted (11-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE          1 (bases 1 to 174913)
JOURNAL        Wilson,S.
COMMENT         requests: clonerequest@sanger.ac.uk
On Jan 13, 2001 this sequence version replaced gi:12044561.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
```

such as compressions and repeats but not necessarily within known annotated repeat sequence elements. Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em.: EMBL; Sw.: SWISSPROT; Tr.: TREMBL; Wp.: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/ChrX> <http://RPI1-382F4> is from the library RPI1-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBACE3.6 This sequence is the entire insert of clone RPI1-382F4.

FEATURES
source

JRES	Location/Qualifiers	repeat_region
Source		
	1. 174913	20958. .22064
	/organism="Homo sapiens"	/note="MER4-internal repeat: matches 1. .1187 of consensus"
	/db_xref="taxon:9606"	22067. .22514
	/chromosome="X"	/note="MER4A2 repeat: matches 1. .504 of consensus"
	/clone="RP11-382F24"	22515. .22657
	/clone_lib="RP11-382F24"	/note="MER4A2 repeat: matches 362. .504 of consensus"
	5849. 5912	22931. .23094
	/note="32 copies 2 mer tt 67% conserved"	/note="MER31-internal repeat: matches 4396. .6198 of consensus"
repeat_region	6418. 6716	23172. .23863
repeat_region	/note="AluJo repeat: matches 1. .290 of consensus"	/note="LTR8 repeat: matches 1. .669 of consensus"
repeat_region	6793. 6859	23866. 25428
repeat_region	/note="MER21B repeat: matches 731. .789 of consensus"	/note="MER4-internal repeat: matches 4491. .6100 of consensus"
repeat_region	6860. 7118	25411. .25599
repeat_region	/note="AluSc repeat: matches 37. .295 of consensus"	/note="MER57-internal repeat: matches 4053. .4244 of consensus"
repeat_region	7119. 7245	25600. 25896
repeat_region	/note="MER21B repeat: matches 574. .731 of consensus"	/note="AluJo repeat: matches 5. .305 of consensus"
repeat_region	7794. 8080	25897. .26883
repeat_region	/note="AluJo repeat: matches 1. .287 of consensus"	/note="MER57-internal repeat: matches 1124. .4053 of consensus"
repeat_region	8640. 8865	27218. 27643
repeat_region	/note="L2 repeat: matches 967. .1225 of consensus"	/note="MER83-internal repeat: matches 1459. .1887 of consensus"
repeat_region	8893. 8991	27806. 28561
repeat_region	/note="L1M4 repeat: matches 4639. .4743 of consensus"	/note="MER57-internal repeat: matches 1304. .2095 of consensus"
repeat_region	9069. 9217	29221. 29355
repeat_region	/note="L1M4 repeat: matches 4751. .4919 of consensus"	/note="LTR8 repeat: matches 546. .691 of consensus"
repeat_region	10536. 10704	29376. 29645
repeat_region	/note="MIR repeat: matches 26. .200 of consensus"	/note="L1PA16 repeat: matches 5891. .6157 of consensus"
repeat_region	10864. 11121	30349. 30999
repeat_region	/note="MER21B repeat: matches 365. .632 of consensus"	/note="L1M2 repeat: matches 5663. .6330 of consensus"
repeat_region	11183. 12011	31000. 31162
repeat_region	/note="L1 repeat: matches 2292. .3139 of consensus"	/note="AluSc repeat: matches 140. .296 of consensus"
repeat_region	12012. 12311	31163. 31451
repeat_region	/note="AluSq repeat: matches 1. .299 of consensus"	/note="AluSc repeat: matches 5. .292 of consensus"
repeat_region	12312. 13246	31452. 31589
repeat_region	/note="L1 repeat: matches 1392. .2292 of consensus"	/note="AluSc repeat: matches 1. .140 of consensus"
repeat_region	13591. 13704	31590. 31844
repeat_region	/note="MERAC repeat: matches 291. .447 of consensus"	/note="L1M2 repeat: matches 5413. .5663 of consensus"
repeat_region	13770. 13848	31834. 31854
repeat_region	/note="MER4A2 repeat: matches 127. .190 of consensus"	/note="L1M2 repeat: matches 5383. .5403 of consensus"
repeat_region	13833. 14176	31855. 32527
repeat_region	/note="MER4A repeat: matches 1. .341 of consensus"	/note="MER51B repeat: matches 1. .617 of consensus"
repeat_region	14193. 16348	32528. 32943
repeat_region	/note="MER4-internal repeat: matches 4087. .6594 of consensus"	/note="L1M2 repeat: matches 5080. .5383 of consensus"
repeat_region	16349. 16649	32945. 33237
repeat_region	/note="AluY repeat: matches 1. .302 of consensus"	/note="AluSc repeat: matches 1. .294 of consensus"
repeat_region	16650. 17319	33238. 33289
repeat_region	/note="MER4-internal repeat: matches 3438. .4087 of consensus"	/note="26 copies 2 mer ta 96% conserved"
repeat_region	17322. 17446	33372. 33395
repeat_region	/note="AluSq/x repeat: matches 13. .138 of consensus"	/note="12 copies 2 mer at 95% conserved"
repeat_region	17447. 17740	33400. 33548
repeat_region	/note="AluSc repeat: matches 1. .291 of consensus"	/note="L1M2 repeat: matches 4927. .5081 of consensus"
repeat_region	17742. 18393	
repeat_region	/note="MER4-internal repeat: matches 2738. .3443 of consensus"	

```

repeat_region 34207..34429
/note="LIM8 repeat: matches 6026..6291 of consensus"
repeat_region 35255..36789
/note="LIM1 repeat: matches 4739..6304 of consensus"
repeat_region 36821..37775
/note="L1 repeat: matches 2350..3318 of consensus"
repeat_region 37795..38632
/note="L1P12 repeat: matches 5324..6159 of consensus"
repeat_region 38633..38934
/note="L1uX repeat: matches 2..303 of consensus"
repeat_region 38935..40052
/note="L1P12 repeat: matches 4189..5324 of consensus"
repeat_region 40314..40494
/note="L1M4 repeat: matches 4619..4814 of consensus"
repeat_region 40643..40952
/note="L1uJ repeat: matches 1..310 of consensus"
repeat_region 41618..41895
/note="L1M3 repeat: matches 152..428 of consensus"
repeat_region 43027..43157
/note="L1M4 repeat: matches 89..231 of consensus"
repeat_region 43212..43271
/note="L1M4 repeat: matches 10..68 of consensus"
repeat_region 43949..44031
/note="L1uJ/FLAM repeat: matches 2..84 of consensus"
repeat_region 44080..44238
/note="LIM8 repeat: matches 6114..6291 of consensus"
repeat_region 45104..45665
/note="L1M4 repeat: matches 3884..4482 of consensus"
repeat_region 45666..45950

Query Match 3.1%; Score 24; DB 9; Length 174913;
Best Local Similarity 100.0%; Pred. No. 0.62;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 679 GTTTAAATGAACACAGCTGAAC 702
|||||
Db 79541 GTTTAAATGAACACAGCTGAAC 79518

RESULT 28
AL450023 Human DNA sequence from clone RP11-552J9 on chromosome X, complete
LOCUS sequence.
DEFINITION AL450023
ACCESSION AL450023.21 GI:16580216
VERSION HTG.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 185532)
AUTHORS Bray-Allen,S.
TITLE Direct Submission
JOURNAL Submitted (02-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk
COMMENT On Nov 1, 2001 this sequence version replaced gi:16214704.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Not that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em., EMBL; Sw.,

```

SWISSPROT; Tr., TREMBL; Wp., WORMPEP: Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/ChrX> RP11-552J9 is from the library RPC1-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACe3.6
This sequence is the entire insert of clone RP11-552J9 The true left end of clone RP13-77011 is at 184976 in this sequence.

```

FEATURES
    source
        1..185532
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="X"
        /clone="RP11-552J9"
        /clone_lib="RPC1-11.2"
        38189..38197
        /note="1329 bases of Tn10 transposon (J01829) removed here. This sequence represents the duplicated flanking sequence of the Tn10."
        86440..86455
        /note="Single clone region. Assembly confirmed by restriction digest data."
        115370..115460
        /note="Single clone region. Reads generated from a transposon library derived from a single pUC clone. Restriction digest data confirm the assembly."

    misc_feature
        51642 a 40741 c 40478 g 52671 t

    misc_feature
        51642 a 40741 c 40478 g 52671 t

BASE COUNT 51642 a 40741 c 40478 g 52671 t
ORIGIN

```

```

Query Match 3.1%; Score 24; DB 9; Length 185532;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 31 CCTGAGGTCTGGATTCTTTCTCC 54
|||||
Db 11271 CCTGAGGTCTGGATTCTTTCTCC 11294

RESULT 29
AX208978 214 bp DNA linear PAT 31-AUG-2001
LOCUS Sequence 818 from Patent WO0157207.
DEFINITION AX208978
ACCESSION AX208978
VERSION AX208978.1 GI:15423401
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 214)
AUTHORS Algate,P.A. and Mannion,J.
TITLE Compositions and methods for the therapy and diagnosis of ovarian cancer
JOURNAL Patent: WO 0157207-A 818 09-AUG-2001;
CORIXA CORPORATION (US)
FEATURES
    source
        1..214
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        74 a 34 c 55 g 51 t

```

```

BASE COUNT 74 a 34 c 55 g 51 t
ORIGIN

Query Match 2.9%; Score 23; DB 6; Length 214;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 645 TAAATGCCAGAACGAGTGAAG 667
|||||

```

Db 122 TAAATGCCAGAGCAGGTGAAG 144

RESULT 30
AX198429
LOCUS AX198429 320 bp DNA linear PAT 29-AUG-2001
DEFINITION Sequence 884 from Patent WO0151513.
ACCESSION AX198429
VERSION AX198429.1 GI:15388750
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 320)
AUTHORS Algate,P.A.
TITLE Ovarian tumor-associated sequences
JOURNAL Patent: WO 0151513-A 884 19-JUL-2001;
CORIXA CORPORATION (US)
FEATURES
source
Location/Qualifiers
1..320
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 96 a 65 c 85 g 74 t
ORIGIN
Query Match 2.9%; Score 23; DB 6; Length 320;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 645 TAAATGCCAGAGCAGGTGAAG 667
|||||
Db 122 TAAATGCCAGAGCAGGTGAAG 144

RESULT 31
AX078298
LOCUS AX078298 580 bp DNA linear PAT 22-FEB-2001
DEFINITION Sequence 102 from Patent WO0107471.
ACCESSION AX078298
VERSION AX078298.1 GI:13157989
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 580)
AUTHORS Hillman,J.L., Lai,P., Tang,Y.T., Yue,H., Au-Young,J., Bandman,O.,
Azimzai,Y., Yang,J., Lu,D.A., Baughn,M.R., Patterson,C. and Shah,P.
TITLE Cell cycle and proliferation proteins
JOURNAL Patent: WO 0107471-A 102 01-FEB-2001;
Incyte Genomics, Inc. (US)
FEATURES
source
Location/Qualifiers
1..580
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 4764233CB1"
BASE COUNT 167 a 135 c 160 g 118 t
ORIGIN
Query Match 2.9%; Score 23; DB 6; Length 580;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 645 TAAATGCCAGAGCAGGTGAAG 667
|||||
Db 520 TAAATGCCAGAGCAGGTGAAG 542

RESULT 32
AX226501
LOCUS AX226501 611 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 7 from Patent WO0161009.

ACCESSION AX226501
VERSION AX226501.1 GI:15555699
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 611)
AUTHORS Malyankar,U.M., Tchernev,V.T., Padigar,M., Taupier,R.J.,
Spytek,K.A., Majumder,K., Guo,X., Spaderna,S.K. and Boldog,F.L.
TITLE Polypeptides and nucleic acids encoding same
JOURNAL Patent: WO 0161009-A 7 23-AUG-2001;
Curagen Corporation (US)
FEATURES
source
Location/Qualifiers
1..611
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 196 a 127 c 152 g 135 t
ORIGIN
Query Match 2.9%; Score 23; DB 6; Length 611;
Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 645 TAAATGCCAGAGCAGGTGAAG 667
|||||
Db 472 TAAATGCCAGAGCAGGTGAAG 494

RESULT 33
HSA318880
LOCUS HSA318880 620 bp mRNA linear PRI 14-JAN-2002
DEFINITION Homo sapiens mRNA for XAGE-2 protein.
ACCESSION AJ318880
VERSION AJ318880.1 GI:18157209
KEYWORDS XAGE-2 gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Zendman,A.J.W., van Kraats,A.A., Weidle,U.H., Ruiters,D.R. and Van
Muljen,G.N.P.
TITLE Expression profile of members of the XAGE cancer/testis antigen
family
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 620)
AUTHORS Zendman,A.J.W.
TITLE Direct Submission
JOURNAL Submitted (14-AUG-2001) Zendman A.J.W., Department of Pathology,
University Medical Center St Radboud, Geert Grooteplein Zuid 24, PO
Box 9101 6500 HB Nijmegen, NETHERLANDS
FEATURES
source
Location/Qualifiers
1..620
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp11.21-Xp11.22"
/tissue_type="Ewing sarcoma"
189..524
/gene="XAGE-2"
189..524
/gene="XAGE-2"
/codon_start=1
/product="XAGE-2 protein"
/protein_id="CAC83007.1"
/db_xref="GI:18157210"
/translation="MSWGRSTYPRPRRSLOPPELGAMLEPTDEEPKEKPTKSR
NPTDQKRDDQGAETQVPLEADLQELCQTKTGDCGEGTDVKGKILPAHEFKMP
EAGEKRSQV"
BASE COUNT 191 a 131 c 156 g 142 t
ORIGIN

Query Match	2.9%	Score 23;	DB 9;	Length 620;	
Best Local Similarity	100.0%	Pred. No. 9.6;			
Matches 23;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	645	TAAATGCCAGAGCAGGTGAAG	667		
Db	485	TAAATGCCAGAGCAGGTGAAG	507		
RESULT 34					
LOCUS	BC009232	648 bp	mRNA	linear	PRI 12-JUL-2001
DEFINITION	Homo sapiens, Similar to G antigen 8, clone MGC:16513				
IMAGE:	3960352, mRNA, complete cds.				
ACCESSION	BC009232				
VERSION	BC009232.1				
KEYWORDS	GI:14328031				
SOURCE	MGC.				
ORGANISM	human.				
REFERENCE	Homo sapiens				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
JOURNAL	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
	1 (bases 1 to 648)				
	Strausberg,R.				
	Direct Submission				
	Submitted (06-JUN-2001) National Institutes of Health, Mammalian				
	Gene Collection (MGC), Cancer Genomics Office, National Cancer				
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,				
	USA				
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov				
COMMENT	Contact: MGC help desk				
	Email: cgabbs-r@mail.nih.gov				
	Tissue Procurement: ATCC				
	cDNA Library Preparation: Rubin Laboratory				
	DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)				
	http://www.systemsbio.org				
	Contact: amadansystemsbiology.org				
	Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia				
	Greene, Mark Kettman and Anuradha Madan				
	Clone distribution: MGC clone distribution information can be found				
	through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov				
	Series: IRAL Plate: 25 Row: m Column: 9				
	This clone was selected for full length sequencing because it				
	passed the following selection criteria: Hexamer frequency ORF				
	analysis, Similarity but not identity to protein.				
FEATURES	Location/Qualifiers				
source	1..648				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="MGC:16513 IMAGE:3960352"				
	/tissue_type="Placenta, Choriocarcinoma"				
	/clone_lib="NIH_MGC_21"				
	/lab_host="DH10B-R"				
	/note="Vector: pOTB7"				
	203..538				
	/codon_start=1				
	/product="Similar to G antigen 8"				
	/protein_id="AAH09232.1"				
	/db_xref="GI:14328032"				
	/translation="MSWRGSEYRPRPSRLOPPELIGAMLEPTDEEPKEKPTKSR				
	NPTDQRDDQAAEIQVPLDADLQELCQTKTGCGCEGTDVKGKILPKAEHFKMP				
	EAGEGKSQV"				
CDS					
BASE COUNT	207 a	134 c	162 g	145 t	
ORIGIN					
Query Match	2.9%	Score 23;	DB 9;	Length 648;	
Best Local Similarity	100.0%	Pred. No. 9.5;			
Matches 23;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	645	TAAATGCCAGAGCAGGTGAAG	667		
Db	499	TAAATGCCAGAGCAGGTGAAG	521		
RESULT 35					
LOCUS	AC098742	220193 bp	DNA	linear	HTG 27-NOV-2001
DEFINITION	Mus musculus chromosome UNK clone RP23-122N2, WORKING DRAFT				
SEQUENCE, 3 unordered pieces.					
AC098742					
AC098742.3	GI:17105319				
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
JOURNAL	1 (bases 1 to 220193)				
	McPherson,J.D. and Waterston,R.H.				
	The sequence of Mus musculus clone				
	Unpublished				
	2 (bases 1 to 220193)				
	McPherson,J.D. and Waterston,R.H.				
	Direct Submission				
	Submitted (31-OCT-2001) Genome Sequencing Center, 4444 Forest Park				
	Parkway, St. Louis, MO 63108, USA				
	On Nov 27, 2001 this sequence version replaced gi:17017630.				
COMMENT					
	----- Genome Center -----				
	Center: Washington University Genome Sequencing Center				
	Center code: WUGSC				
	Web site: http://genome.wustl.edu/gsc/index.shtml				
	Contact: submissions@watson.wustl.edu				
	----- Project Information -----				
	Center project name: M_BA0122N02				
	----- Summary Statistics -----				
	Sequencing vector: M13; 84%				
	Chemistry: Dye-primer ET; 0% of reads				
	Chemistry: Dye-terminator Big Dye; 100% of reads				
	Assembly program: Phrap; version 0.990319				
	Consensus quality: 219149 bases at least Q40				
	Consensus quality: 219343 bases at least Q30				
	Consensus quality: 219466 bases at least Q20				
	Insert size: 90ul; agarose-1p				
	Insert size: 222856; sum-of-contigs				
	Quality coverage: 13.57 in Q20 bases; agarose-1p				
	Quality coverage: 13.91 in Q20 bases; sum-of-contigs				

	* NOTE: This is a 'working draft' sequence. It currently				
	* consists of 3 contigs. The true order of the pieces				
	* is not known and their order in this sequence record is				
	* arbitrary. Gaps between the contigs are represented as				
	* runs of N, but the exact sizes of the gaps are unknown.				
	* This record will be updated with the finished sequence				
	* as soon as it is available and the accession number will				
	* be preserved.				
	* 1 218685: contig of 218685 bp in length				
	* 218686 218785: gap of unknown length				
	* 218786 219435: contig of 650 bp in length				
	* 219436 219535: gap of unknown length				
	* 219536 220193: contig of 658 bp in length.				
FEATURES	Location/Qualifiers				
source	1..220193				
	/organism="Mus musculus"				
	/db_xref="taxon:10090"				
	/chromosome="UNK"				
	/clone="RP23-122N2"				
	/note="assembly_name:Contig65"				
misc_feature	1..218685				
	/note="assembly_name:Contig65"				
misc_feature	218786..219435				
	/note="assembly_name:Contig20"				
misc_feature	219536..220193				
	/note="assembly_name:Contig33"				
BASE COUNT	70571 a	40361 c	40477 g	68584 t	200 others

ORIGIN	Query Match 2.9%; Score 23; DB 2; Length 220193;									
	Best Local Similarity 100.0%; Pred. No. 1.7;									
QY	Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
	253 GAGCAAGGAGGCCGAGGAGTGG 275									
Db	200939 GAGCAAGGAGGCCGAGGAGTGG 200961									
RESULT 36	CEC33G3/c									
	LOCUS									
DEFINITION	Caenorhabditis elegans cosmid C33G3, complete sequence.									
	278540									
ACCESSION	278540.1 GI:1507775									
	HTG: Beta-mannosidase; Lys-tRNA; Transfer RNA.									
KEYWORDS	Caenorhabditis elegans.									
	Caenorhabditis elegans.									
SOURCE	Rhabditidae; Metazoa; Nematoda; Chromadorea; Rhabditida;									
	Rhabditidae; Rhabditidae; Peleoderinae; Caenorhabditis.									
REFERENCE	1 (sites)									
	none.									
AUTHORS	Genome sequence of the nematode C. elegans: a platform for									
	investigating biology. The C. elegans Sequencing Consortium									
TITLE	Science 282 (5396), 2012-2018 (1998)									
	99069613									
JOURNAL	The C.elegans Sequencing Consortium.									
	Matthews, L.									
MEDLINE	Direct Submission									
	Submitted (23-AUG-1996) Nematode Sequencing Project, Sanger									
REMARK	Institute, Hinxton, Cambridge CB10 1SA, England and department of									
	Genetics, Washington University, St. Louis, MO 63110, USA. E-mail:									
AUTHORS	jes@sanger.ac.uk or rw@nematode.wustl.edu									
	Coding sequences below are predicted from computer analysis, using									
TITLE	Predictions from Genefinder (P. Green, U. Washington), and other									
	available information.									
JOURNAL	Current sequence finishing criteria for the C. elegans genome									
	sequencing consortium are that all bases are either sequenced									
MEDLINE	unambiguously on both strands, or on a single strand with both a									
	dye primer and dye terminator reaction, from distinct subclones.									
REMARK	Exceptions are indicated by an explicit note.									
	IMPORTANT: This sequence is not the entire insert of clone C33G3.									
AUTHORS	It may be shorter because we only sequence overlapping sections									
	once, or longer because we arrange for a small overlap between									
TITLE	neighbouring submissions.									
	The true left end of clone C33G3 is at 1 in this sequence. The true									
JOURNAL	right end of clone C33G3 is at 9942 in									
	sequence Z81064.									
REMARK	The true left end of clone F16B12 is at 33555 in this sequence. The									
	true right end of clone C04C11 is at 3475 in this sequence. The									
AUTHORS	start of this sequence (1. 104) overlaps with the end of sequence									
	272501.									
TITLE	The end of this sequence (33555. .33658) overlaps with the start of									
	sequence Z81064.									
JOURNAL	For a graphical representation of this sequence and its analysis									
	see:- http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?									
FEATURES	name=C33G3									
	IMPORTANT: This sequence is NOT necessarily the entire insert of									
source	the specified clone. It may be shorter because we only sequence									
	overlapping sections once, or longer because we arrange for a small									
location/Qualifiers	overlap between neighbouring submissions.									
	1. .33658									
/organism="Caenorhabditis elegans"	/db_xref="taxon:6239"									
	/chromosome="x"									
/clone="C33G3"	join(3893..4063,5154..5198,5368..5434,5705..5850,									
	7934..7084,7200..7334,7794..7946,7999..8247,8388..8884,									
gene	9084..9303,9350..9445,9581..9780,11849..11960)									
	translation="MKEIQFNSSETSDLMQHEVNTLEQCLYYLAGESAKKLAKINIP									
/translation="MKEIQFNSSETSDLMQHEVNTLEQCLYYLAGESAKKLAKINIP	HAIASIPNYSQQPVTINFAQFTLOHFMELKDISIHGYLELSHWHDRLMNQDT									
	WKKNKLVVSHFHHVNPVLLGQSNENHLKNGDAFEIRKRVETTNQNOYSKAVFSLRTE									

/gene="C33G3.1"
join(3893..4063,5154..5198,5368..5434,5705..5850,
6794..7084,7200..7334,7794..7946,7999..8247,8388..8884,
9084..9303,9350..9445,9581..9780,11849..11960)
/gene="C33G3.1"
/note="contains similarity to Pfam domain: PF00640
(Phosphotyrosine interaction domain (PTB/PID).),
Score=21.1, E-value=3.5e-05, N=1
cDNA EST yk259a5.5 comes from this gene"
/codon_start=1
/protein_id="CAB01738.2"
/db_xref="GI:6434268"
/db_xref="SPTREMBL:Q93325"
/translation="MIRFDPDESSESEKRFYYLRMLAKPGKPIKTKFTVTSRDED
VASASNTSDQLESQAMRVVRTIGOAFVCHVAQDMOEKHEDEAKSKISQSEDE
AGNALDVIIEGRGREDSSSPMEAPPVGGPLKRLSLFOPRKPSTTSSSGGTAI
DTTAIPENLIEPTNSHTIOPRAPELVQLOPOTALPYOOKPQSLNLLQOQFNTLP
SOMPSTQTLPSLSENGQSHIPRMTMPPNPYPATLPHRTWAQDLPQHSQSTL
EQNYMYQPQMLPSSSLPFLCLSPVMVSPYATLQNMQQSLDQPHQSDITLL
DOYNLMRSQDQAQSVQVAGCVOLLRDQLTSETTARLEAQSRTHLLSANRDL
PGSSIRLNYQVOPLADLRAGSLPPVKERRKDEGTRTEPESNAEDTTIDYSSSD
QYRTSNVMPKSFENILSNPLVDINVPSCAAMSRMEQFDMGDTGCTSTPPKKEKK
PSSGILGERSRMSFNPGLKREKEROQQLMEFEDTLEDDSPRSIPSPSPKARNITL
DSLFPKQDDPTIADREPQQLPPQSQNOKKNTAVNLLMPTMPASSSLVTMYPPMR
QPAVPVNIQPKVDVFRKTKTLSDMDIAEEPSEMDPNRNNILPSSNTSSMKRRGL
PPNTDVEIKIEDYINRNVDRSKLPTSLTLRLTQQAQDNLGSLNPGYQ"
10280..10354
/gene="C33G3.t1"
/note="CTT Lys K-tRNA
predicted using tRNAscan-SE-1.11
preliminary prediction
similar to tRNA-Lys"
/product="tRNA-Lys"
10280..10354
/gene="C33G3.t1"
complement(11047..11121)
/gene="C33G3.t3"
/note="CTT Lys K-tRNA
predicted using tRNAscan-SE-1.11
preliminary prediction
similar to tRNA-Lys"
/product="tRNA-Lys"
complement(11047..11121)
/gene="C33G3.t3"
complement(12554..12626)
/gene="C33G3.t2"
/note="TTT Lys K-tRNA
predicted using tRNAscan-SE-1.11
preliminary prediction
similar to tRNA-Lys"
/product="tRNA-Lys"
complement(12554..12626)
/gene="C33G3.t2"
complement(13042..13243,13286..13400,13455..13560,
13607..13855,15236..15419,15543..15694,15803..15905,
16167..16285)
/gene="C33G3.3"
complement(join(13042..13243,13286..13400,13455..13560,
13607..13855,15236..15419,15543..15694,15803..15905,
16167..16285))
/gene="C33G3.3"
/note="predicted using Genefinder
Weak similarity to acetylcholine receptor alpha-7 chain
(SW:ACH7_HUMAN)
cDNA EST EMBL:Z14399 comes from this gene"
/codon_start=1
/protein_id="CAB01736.1"
/db_xref="GI:387464"
/db_xref="SPTREMBL:Q93323"
/translation="MKEIQFNSSETSDLMQHEVNTLEQCLYYLAGESAKKLAKINIP
HAIASIPNYSQQPVTINFAQFTLOHFMELKDISIHGYLELSHWHDRLMNQDT
WKKNKLVVSHFHHVNPVLLGQSNENHLKNGDAFEIRKRVETTNQNOYSKAVFSLRTE


```

CDDTDFENYNDVYKCCFSEFPQDREVIOFTSSGLPIFTDPKNFRDYGWGVSGCTVPE
SFDPSEIAQFCNLKRAHSLKIELAIFLFTTALFLPPIFGSVKIQIYLKMEV
MGLQMTLLIFSTIAPLSTASTPKPMRELAIALVFNLIISITTSIIIFCCMOVKRT
LPPGRVTFQANFINGELGVNLISGVEYNLDKYDEQVSONSYQKDWANVFRAAHAVL
MGTISAIVIFVIYVCL"
Join(19807)..19939,20519..20670,20805..21005,21238..21314,
21363..21496,21547..21801,21847..22151,22845..23146,
23258..23401,23445..23568,23613..23930,23987..24168,
24215..24329,24544..24804)
/genes="C33G3.4"
Join(19807)..19939,20519..20670,20805..21005,21238..21314,
21363..21496,21547..21801,21847..22151,22845..23146,
23258..23401,23445..23568,23613..23930,23987..24168,
24215..24329,24544..24804)
/genes="C33G3.4"
/note="predicted using Genefinder
Similarity to Bovine beta-mannosidase (TR:G704360)
cDNA EST Yk23h3.3 comes from this gene
cDNA EST Yk26d5.3 comes from this gene
cDNA EST Yk23h3.5 comes from this gene
cDNA EST Yk26d5.5 comes from this gene
cDNA EST Yk44e10.5 comes from this gene
cDNA EST Yk128c10.3 comes from this gene
cDNA EST Yk134e12.5 comes from this gene
cDNA EST Yk159c11.5 comes from this gene
cDNA EST Yk177c7.5 comes from this gene
cDNA EST Yk159c11.3 comes from this gene
cDNA EST Yk177c7.3 comes from this gene
cDNA EST Yk476d8.3 comes from this gene
cDNA EST Yk476d8.5 comes from this gene
cDNA EST Yk415g9.3 comes from this gene
cDNA EST Yk415g9.5 comes from this gene"
/codon_start=1
/protein_id="CAB01737.1"
/db_xref="GI:3874645"
/db_xref="SWISS-PROT:Q93324"
/translation="MRTSLVGLFLLFQLHTTHGYNTLVNLGNWFESSSKNTVNGT
GVPGDIVSDLYASGITIDNPLFGENHLNKLWIAEDDWTYSRRKFLIDDDTVGAFLEI
ESVDITATVYNGVKLVHGNQFLPYHNVTDI IALGENDITIKFKSVKYAEKRADE
YKIFGHSLPDCNDGLVHGCHONFIRKAOYSFANDGSPFTVCIPISTLTINIRG
QIFHDNMTFAHGKWKVAFEDTFHYGARTIEYSVQIPELGIKESDYRLSATKSL
QTRSKNLSLIPMEHEPERMPNGSQKLYDVVSMGQGVKKKIKFKTVELVODL
IDPKPKGRNRYFKIINDEPVLKGTNIPVSMFRSDRENIAKTEFLDSDVAEVMGNA
IRVWGGFVSHNFYAYKKGILVMQDLMFACALYPTTEF IQNAEEVSYNVDRIIS
QHTVSYSGNNEAAIARGHMKASNTESQOVKDYLLYORLAKIAKVAPTIPFI
MSPSGNTEEGGVSKNPYDVRGYDIIHYNEFVNLNRDDTYLTPRCASEYGVQSY
MKETLWNINSDWEYTSKAMFHROHPGGIATNLLMIFQHLPIPAECGSKSVQVPS
CKTISASMSRLATFSQVHQSIKLTQTLTRFRNTTNEGILNRCAMTQWLDV
WAAPTWTSTIDFQNMKAHYEARFNSVAVSFADETDFNLKVFLLNDNPYLLHNT
VNVQMLSGNGLDPILTNEPHIDSPVAGSSEVLKTCITFSKITELSYLYSTLYDSS
GVKIHEDVLVPLFPEVDNFTGQVQISDVQRIDEKTYHLTITTDVRSFPFWITCKKP
FTGWFSDNGFHMQLRLKIRLILAEVLEKSDFTVCNINKYV"
complement(Join(24924..25150,25205..25435,26181..26264,
26308..26916,27080..27173,27220..27282))
/genes="C33G3.5"
complement(Join(24924..25150,25205..25435,26181..26264,
26308..26916,27080..27173,27220..27282))
/genes="C33G3.5"
/note="predicted using Genefinder"
/codon_start=1
/protein_id="CAB01735.1"
/db_xref="GI:3874643"
/db_xref="SPTREMBL:Q93322"
/translation="MIIPVTWVVFHLLFLCLCHAEYYDYDATIRQVPIYDGSKFYE
MSPSENAFTLNGLLYSTENVSIVSHVSPPEHFRSHSTDEDDLDLVMXYKDSHF
DKTGQLRYASVPVIEDLVNDILNVIFGRKEYIRLLNPNRAEMTDIVREMVRVPTQE
VLKQQLFHKILASRAYEYKFFKPMLEDNNRQKLLFEGRHLLDSHNGTLSTSDTLEIN
2.8%; Score 22; DB 3; Length 33658;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match
Best Local Similarity 100.0%; Pred. No. 83; Length 581;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 26926 AGTTTTCAGCTTTCACCAAAA 26905
|||||
RESULT 37
AF143879 Homo sapiens clone IMAGE:120631 mRNA linear PRI 27-MAY-1999
LOCUS AF143879 581 bp mRNA
DEFINITION Homo sapiens clone IMAGE:120631 mRNA sequence.
ACCESSION AF143879
VERSION AF143879.1 GI:4895021
KEYWORDS FLI_CDNA.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Persson,A.E., Lundeborg,J. and Uhlen,M.
TITLE EU-IMAGE: Full-insert length sequencing of human cDNA clones
JOURNAL Unpublished
REFERENCE
AUTHORS Persson,A.E., Lundeborg,J. and Uhlen,M.
TITLE Direct Submission
JOURNAL Submitted (19-APR-1999) Biochemistry and Microbiology,
Biotechnology, Royal Institute of Technology, Teknikringen 34,
Stockholm S-100 44, Sweden
FEATURES
Source
1..581
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:120631"
BASE COUNT 154 a 135 c 128 g 164 t
ORIGIN

Query Match
Best Local Similarity 2.7%; Score 21; DB 9; Length 581;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 761 GCTTTCACCAAAAAA 781
|||||
Db 553 GCTTTCACCAAAAAA 573
|||||

RESULT 38
HS5UL103HG 1665 bp DNA linear VRL 02-MAR-1995
LOCUS HS5UL103HG 1665 bp DNA
DEFINITION Murine cytomegalovirus UL103 homologue gene, complete cds, UL102
and UL104 homologue genes, partial cds.
ACCESSION L29246
VERSION L29246.1 GI:468306
KEYWORDS Murine cytomegalovirus (strain K181) DNA.
SOURCE Mouse cytomegalovirus 1
ORGANISM Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Betaherpesvirinae; Muromegalovirus.
REFERENCE
AUTHORS Lyons,P.A., Dallas,B., Carrello,C., Shellam,G.R. and Scalzo,A.A.
TITLE Mapping and transcriptional analysis of the murine cytomegalovirus
homologue of the human cytomegalovirus UL103 open reading frame
JOURNAL Virology 204 (2), 835-839 (1994)
MEDLINE 95027718
FEATURES
Source
1..1665
/organism="Mouse cytomegalovirus 1"
/strain="K181"
/db_xref="taxon:10366"
<1..177
/note="UL102 homologue"
/codon_start=1
/protein_id="AAA62598.1"
/db_xref="GI:688449"
/translation="ETTPGTANYRQPKQKTIPTTTPIGLPPSPSPAPEVSPRFRSPQPK
LETLDRNRVQHLNG"
155..1108

CDS

CDS
```

```
/note="UL103 homologue"
/codon_start=1
/protein_id="AAA62599.1"
/db_xref="GI:468307"
/translation="MSSTLMVGMVHNDYVRLMAPEIVSITVSDROIWPHTEGG
ELINQAQRADSCRSFLEGCFLFDALSDALSELPLSSIRTKHRAVFRPKNTIDF
TLCMLLFAIESPLUSRGLHLEHVLFTGTRPTGLRMICKSCIKLVCTSLYLFFDET
DPTRTRVPTCMYKQTRQAQALAEATYFGVODISSMSLSVLTLDRTVKOGDVG
DLAAEVLTNCNVFYVPLNGSVVTRWLSNNDNPKRMYPTDTTDTTIPPTPTK
KNMTTATSHIALLSNTNTTTTTPGADTKKSNKK"
complement(1256..>1665)
/note="UL104 homologue"
/codon_start=3
/protein_id="AAA62600.1"
/db_xref="GI:688450"
/translation="GRMVAVOPLDGLHNDICYNKYKTFEVLRLVLEKVIIVLSOD
QNTAVNNSGTGEIDVAIKQEVTDDAIVSDTTGSDGLDAALRALRETTANTVGMVVE
RTDTHNLTEFNDFVQOTMDYFVERYSAFSTMN"
BASE COUNT 353 a 513 c 459 g 340 t
ORIGIN
```

```
Query Match 2.7%; Score 21; DB 14; Length 1665;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 551 TGATCTGCAAGAGCTGCATCA 571
|||||
Db 573 TGATCTGCAAGAGCTGCATCA 593
```

```
RESULT 39
AC013875/c
LOCUS
DEFINITION
AC013875
AC013875.1 GI:6437460
VERSION
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 9971)
Adams, M. and Venter, J.C.
Direct Submission
Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
This sequence was identified as CDM:10213220 by the submitter.
For further information on this sequence e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
Source
1..9971
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
BASE COUNT 2516 a 2453 c 2501 g 2501 t
ORIGIN
```

```
Query Match 2.7%; Score 21; DB 2; Length 9971;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 ACACACACAAACACAGAACCA 413
|||||
Db 9575 ACACACACAAACACAGAACCA 9555
```

```
RESULT 40
AC073266/c
LOCUS
```

```
DEFINITION Homo sapiens chromosome 19 clone RP11-393E18, WORKING DRAFT
SEQUENCE, 30 unordered pieces.
AC073266
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Waterston, R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 200688)
Waterston, R.H.
Direct Submission
Submitted (12-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Aug 11, 2000 this sequence version replaced gi:8920670.
COMMENT
```

```
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0393E18
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 182974 bases at least Q40
Consensus quality: 187646 bases at least Q30
Consensus quality: 190053 bases at least Q20
Insert size: 168000; agarose-fp
Quality coverage: 5.50 in Q20 bases; agarose-fp
Quality coverage: 4.79 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 30 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
```

```
1 1456: contig of 1456 bp in length
* 1457 1556: gap of unknown length
* 1557 3865: contig of 2309 bp in length
* 3866 3965: gap of unknown length
* 3966 5298: contig of 1333 bp in length
* 5299 5398: gap of unknown length
* 5399 7350: contig of 1952 bp in length
* 7351 7450: gap of unknown length
* 7451 8753: contig of 1303 bp in length
* 8754 8854: gap of unknown length
* 8854 10542: contig of 1688 bp in length
* 10542 12310: contig of 1669 bp in length
* 12311 12410: gap of unknown length
* 12411 14524: contig of 2114 bp in length
* 14525 14624: gap of unknown length
* 14625 16559: contig of 1934 bp in length
* 16559 16659: gap of unknown length
* 16659 18499: contig of 1840 bp in length
* 18499 18599: gap of unknown length
* 18599 21245: contig of 2646 bp in length
* 21245 21345: gap of unknown length
* 21345 23978: contig of 2634 bp in length
* 23979 24078: gap of unknown length
```

```
* 24079 27768: contig of 3690 bp in length
* 27769 27868: gap of unknown length
* 27869 31566: contig of 3798 bp in length
* 31566 31766: gap of unknown length
* 31766 36499: contig of 4733 bp in length
* 36499 44328: gap of unknown length
* 44328 44329: contig of 7729 bp in length
* 44329 48785: gap of unknown length
* 48785 48886: gap of unknown length
* 48886 53783: contig of 4797 bp in length
* 53783 60899: gap of unknown length
* 60899 69528: contig of 7016 bp in length
* 69528 78414: gap of unknown length
* 78414 87329: gap of unknown length
* 87329 95501: contig of 8630 bp in length
* 95501 105284: gap of unknown length
* 105284 115544: gap of unknown length
* 115544 127535: gap of unknown length
* 127535 137337: gap of unknown length
* 137337 156915: gap of unknown length
* 156915 173311: gap of unknown length
* 173311 200688: contig of 16296 bp in length
* 200688 27277: gap of unknown length
* 27277 27277: contig of 27277 bp in length.

FEATURES
source      Location/Qualifiers
1..200688
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/clone="RP11-393E18"
1..1456
/note="assembly_name:Contig9"
1537..3865
/note="assembly_name:Contig11"
3966..5298
/note="assembly_name:Contig12"
5399..7350
/note="assembly_name:Contig13"
7451..8753
/note="assembly_name:Contig14"
8854..10541
/note="assembly_name:Contig15"
10642..12310
/note="assembly_name:Contig16"
12411..14524
/note="assembly_name:Contig17"
14625..16558
/note="assembly_name:Contig18"
clone_end:r7
vector_side:left"
16659..18498
/note="assembly_name:Contig19"
18599..21244
/note="assembly_name:Contig20"
21345..23978
/note="assembly_name:Contig21"
24079..27768
/note="assembly_name:Contig22"
27869..31666
/note="assembly_name:Contig23"
31767..36499
/note="assembly_name:Contig24"
```

```
misc_feature 36600..44328
/note="assembly_name:Contig25"
misc_feature 44429..48785
/note="assembly_name:Contig26"
clone_end:SP6
vector_side:right"
48886..53682
/note="assembly_name:Contig27"
53783..60798
/note="assembly_name:Contig28"
60899..69528
/note="assembly_name:Contig29"
69629..78413
/note="assembly_name:Contig30"
78514..87228
/note="assembly_name:Contig31"
87329..95400
/note="assembly_name:Contig32"
95501..105284
/note="assembly_name:Contig33"
105385..115544
/note="assembly_name:Contig34"
115645..127534
/note="assembly_name:Contig35"
127635..137337
/note="assembly_name:Contig36"
137438..156915
/note="assembly_name:Contig37"
157016..173311
/note="assembly_name:Contig38"
173412..200688
/note="assembly_name:Contig39"

BASE COUNT 60199 a 40808 c 41989 g 54792 t 2900 others
ORIGIN

Query Match      2.7%:  Score 21;  DB 2;  Length 200688;
Best Local Similarity 100.0%;  Pred. No. 15;
Matches 21;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY  393 ACACACACAAACACACAGAACCA 413
    |||||.....|
DB 69920 ACACACACAAACACACAGAACCA 69900

Search completed: October 16, 2002, 22:51:31
Job time : 2917 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: October 16, 2002, 22:51:36 ; Search time 1856 Seconds
(without alignments)
8805.832 Million cell updates/sec

Title: US-09-658-824-808
Perfect score: 781
Sequence: 1 gggcgaggctgtgagccgg.....ctttcaccaaaaaaaaaa 781

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.roi.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
------------	-------	-------	-------	--------	----	----	-------------

1	781	100.0	781	6	AX063181
2	740	94.8	740	9	HSA318878
3	526	67.3	529	9	BC009538
4	392	50.2	6060	9	HSA400997
c	392	50.2	180859	2	AC025553
6	391	50.1	396	6	AX245237
7	391	50.1	626	6	AF251237
8	390.6	50.0	399	6	AX062442
9	387	49.6	463	6	HSB280447
10	369	47.2	457	6	AX063163
11	363	46.5	481	9	HSB318879
12	354	45.3	479	6	AX062443
13	350	44.8	461	6	AX063165
14	305	39.1	391	6	AX246260
15	214.4	27.5	493	9	HSB318881
16	213.2	27.3	648	9	BC009232
17	212.8	27.2	620	9	HSB318880
18	205.8	26.4	81953	2	AC018835
19	202	25.9	114344	2	AL445227
c	202	25.9	149749	9	AL445236
21	201.8	25.8	611	6	AX226501
22	200.4	25.7	93419	2	AC002415
23	182	23.3	185532	9	AL450023
c	174	22.3	119024	9	AL159987
25	169	21.6	580	6	AX078298
c	167.8	21.5	93419	2	AC002415
c	167.8	21.5	114344	2	AL445227
28	167.8	21.5	149749	9	AL445236
29	166.4	21.3	130005	9	AL590240
c	164.2	21.0	20587	9	HSB18556A
c	164.2	21.0	62493	9	HSB193G15
c	164.2	21.0	62493	9	HSB193G15
c	135	17.3	320	6	AX198429
35	133.2	17.1	214	6	AX208978
36	122.4	15.7	164591	2	AC090022
37	113.4	14.5	109469	9	AC004535
38	110	14.1	528	9	AF055473
39	110	14.1	539	6	AR028492
40	110	14.1	539	6	I55855
41	110	14.1	540	6	AR028490
42	110	14.1	540	6	I55853
43	110	14.1	551	9	BC018052
44	108.4	13.9	532	6	AR028491
45	108.4	13.9	532	6	I55854

ALIGNMENTS

RESULT 1
AX063181
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE

AX063181
Sequence 808 from Patent WO0100828.
AX063181
AX063181.1

GI:12541027
human.
Homo sapiens

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

Source

Location/Qualifiers

1..781

/organism="Homo sapiens"

/db_xref="taxon:9606"

230 a 181 c 234 g 136 t

BASE COUNT

AX063181 Sequence
AJ318878 Homo sapi
BC009538 Homo sapi
AJ400997 Homo sapi
AC025553 Homo sapi
AX245237 Sequence
AF251237 Homo sapi
AX062442 Sequence
AJ290447 Homo sapi
AX063163 Sequence
AJ318879 Homo sapi
AX062443 Sequence
AX063165 Sequence
AX246260 Sequence
AJ318881 Homo sapi
BC009232 Homo sapi
AJ318880 Homo sapi
AC018835 Homo sapi
AL445227 Homo sapi
AL445236 Human DNA
AX226501 Sequence
AC002415 Homo sapi
AL450023 Human DNA
AL159987 Human DNA
AX078298 Sequence
AC002415 Homo sapi
AL445227 Homo sapi
AL445236 Human DNA
AL590240 Human DNA
268274 Human DNA s
AL117391 Human DNA
AL096938 Human DNA
AL645949 Homo sapi
AX198429 Sequence
AX208978 Sequence
AC090022 Homo sapi
AC004535 Homo sapi
AF055473 Homo sapi
AR028492 Sequence
I55855 Sequence 18
AR028490 Sequence
I55853 Sequence 16
BC018052 Homo sapi
AR028491 Sequence
I55854 Sequence 17

AX063181
Sequence 808 from Patent WO0100828.
AX063181
AX063181.1

781 bp
DNA
linear
PAT 24-JAN-2001

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 781)
Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedwick,T.S.,
Carter,D., Retter,M.W. and Mannion,J.
Compositions and methods for the therapy and diagnosis of lung
cancer
Patent: WO 0100828-A 808 04-JAN-2001;
CORIXA CORPORATION (US)

Location/Qualifiers
1..781
/organism="Homo sapiens"
/db_xref="taxon:9606"

230 a 181 c 234 g 136 t

BASE COUNT

Db	481	CAGCTGAGATCCCACTGCGCGACATGGAAGGTGATCTCGAAGAGCTGCATCAGTCAACA	540
Qy	580	CCGGGGATAAATCTGGATTTCGGTTCGGCGCTCAAGGTGAAGATAATACTAAACAGCAA	639
Db	541	CCGGGGATAAATCTGGATTTCGGTTCGGCGCTCAAGGTGAAGATAATACTAAACAGCAA	600
Qy	640	CACGTGAAAATGCCAAGACGAGGTGAACAGCAACACCAAGTTTAATCAAGACAAAGCTGA	699
Db	601	CACGTGAAAATGCCAAGACGAGGTGAACAGCAACACCAAGTTTAATCAAGACAAAGCTGA	660
Qy	700	AAACAACGCAAGCTGGTTTTATATTAGATATTGACTTAAACTATCTCAATAAAAGTTTTC	759
Db	661	AAACAACGCAAGCTGGTTTTATATTAGATATTGACTTAAACTATCTCAATAAAAGTTTTC	720
Qy	760	AGCTTTCACCAAAAAA 779	
Db	721	AGCTTTCACCAAAAAA 740	
RESULT 3			
LOCUS	BC009538	529 bp mRNA linear	PK1 22-OCT-2001
DEFINITION	Homo sapiens, clone MGC:10063 IMAGE:3893227, mRNA, complete cds.		
ACCESSION	BC009538		
VERSION	BC009538.1	GI:16306934	
KEYWORDS	MGC.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Strausberg,R.		
JOURNAL	Submitted (29-JUN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: DCTD/DRP/Gazdar cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www-shgc.stanford.edu Contact: (Dickson, Mark) mcd@paxil.stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.		
FEATURES	source	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAK Plate: 14 Row: g Column: 8. Location/Qualifiers 1..529 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="MGC:10063 IMAGE:3893227" /tissue_type="Lung, carcinoma, large cell undifferentiated." /clone_lib="NIH_MGC_69" /lab_host="DH10B" /note="Vector: pCMV-SPORT6" 184..429 /codon_start=1 /product="Unknown (protein for MGC:10063)" /protein_id="AAH09538.1" /db_xref="GI:16306935" /translation="MESPKKKNOQLKVGILHLSGRKKIRIQLRSQATWVKVICKSCI SOTPGINLDLGSVGVKVIIPKEEHCKMPEAGEEPQV"	
BASE COUNT	184 a	116 c	146 g 83 t
ORIGIN	exon		

Query Match	67.3%	Score 526;	DB 9;	Length 529;
Best Local Similarity	100.0%;	Pred. No. 8e-134;		
Matches 526;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	256	GAAGAGAGGCGCAGAGGTGGAGGGCTTCAGCGAAGCTGTGGGGGTATCCG	315	
Db	1	GAAGAGGCGCGAGGTGGAGGGCTTCAGCGAAGCTGTGGGGGTATCCG	60	
Qy	316	AGTCCCAAGACCTTGAACCCCGACAGAAAGATTCTGGACTCCCGACCGGACCAAGA	375	
Db	61	AGTCCCAAGACCTTGAACCCCGACAGAAAGATTCTGGACTCCCGACCGGACCAAGA	120	
Qy	376	GAGGACCGGCATGAGCGACACACAAACACAGAACCAACAGCCAGTCCCGAGGACCCA	435	
Db	121	GAGGACCGGCATGAGCGACACACAAACACAGAACCAACAGCCAGTCCCGAGGACCCA	180	
Qy	436	GTAATGGAGAGCCCAAAAAGAACAGCAGCTGAAAGTCGGGATCCTACACCTGGGC	495	
Db	181	GTAATGGAGAGCCCAAAAAGAACAGCAGCTGAAAGTCGGGATCCTACACCTGGGC	240	
Qy	496	AGCAGACAGAAGAAGATCAGGATACAGTCCAGTCCAGTCCCGGACATGGAAGTGATC	555	
Db	241	AGCAGACAGAAGAAGATCAGGATACAGTCCAGTCCAGTCCCGGACATGGAAGTGATC	300	
Qy	556	TGCAAGAGCTGCATCAGTCAACACCCGGGGATAAACTCTGGATTTCGGGCGGTCAAG	615	
Db	301	TGCAAGAGCTGCATCAGTCAACACCCGGGGATAAACTCTGGATTTCGGGCGGTCAAG	360	
Qy	616	GTGAAGATAATACCTAAAGAGGAACACTGTTAAATGCCAGAGCAGTGAAGACAACCA	675	
Db	361	GTGAAGATAATACCTAAAGAGGAACACTGTTAAATGCCAGAGCAGTGAAGACAACCA	420	
Qy	676	CAAGTTTAAATGAAGACAAGCTGAAACAACCAAGCTGGTTTATATTAGATATTGACT	735	
Db	421	CAAGTTTAAATGAAGACAAGCTGAAACAACCAAGCTGGTTTATATTAGATATTGACT	480	
Qy	736	TAACTATCTCAATAAAGTTTTCAGCTTTCACCAAAAAA 781		
Db	481	TAACTATCTCAATAAAGTTTTCAGCTTTCACCAAAAAA 526		
RESULT 4				
LOCUS	HSA400997	6060 bp DNA linear	PRI 05-MAY-2001	
DEFINITION	Homo sapiens XAGE-1b gene for cancer/testis-associated protein XAGE-1b, exons 1-4.			
ACCESSION	AJ400997			
VERSION	AJ400997.1	GI:13992557		
KEYWORDS	cancer/testis-associated protein; XAGE-1b gene.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	Zendman,A.J.W., van Kraats,A.A., Weidle,U.H., Rutter,D.J. and Van Muijen,G.N.P.			
TITLE	XAGE-1b, a novel cancer/testis-associated gene, induced in melanoma metastases			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 6060)			
AUTHORS	zendman,A.J.W.			
TITLE	Direct Submission			
JOURNAL	Submitted (24-MAY-2000) Zendman A.J.W., Pathology, University Medical Center St. Radboud, Geert Grooteplein Zuid 24, Nijmegen, 6500 HB POBox 9101, NETHERLANDS			
FEATURES	source	Location/Qualifiers 1..6060 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="X" /map="Xp11.21-Xp11.22" /tissue_type="melanoma metastasis" 659..736		
BASE COUNT	184 a	116 c	146 g 83 t	
ORIGIN	exon			

gene	/gene="XAGE-1b"									
	/number=1									
	659. .5636									
	/gene="XAGE-1b"									
	join(659. .736,2143. .2284,2812. .2938,5529. .5636)									
mRNA	/gene="XAGE-1b"									
	737. .2142									
intron	/gene="XAGE-1b"									
	/number=1									
exon	2143. .2284									
	/gene="XAGE-1b"									
CDS	/number=2									
	join(2189. .2284,2812. .2938,5529. .5551)									
intron	/gene="XAGE-1b"									
	/codon_start=1									
	/product="cancer/testis-associated protein XAGE-1b"									
	/protein_id="CAC38108.1"									
	/db_xref="GI:13992558"									
exon	/translation="MESPKKNOQLKVGILHLSRKKIRIOLRSQCATWKVICKSCI									
	SOTPGINLDLGSQVKVIIPKEEHKMPAGEEQPV"									
	2285. .2811									
	/gene="XAGE-1b"									
	/number=2									
intron	2812. .2938									
	/gene="XAGE-1b"									
exon	/number=3									
	2939. .5528									
intron	/gene="XAGE-1b"									
	/number=3									
exon	5529. .5636									
	/gene="XAGE-1b"									
polya_signal	/number=4									
	5614. .5619									
BASE COUNT	/gene="XAGE-1b"									
	1700 a 1184 c 1420 g 1756 t									
ORIGIN										
Query Match 50.2%; Score 392; DB 9; Length 6060;										
Best Local Similarity 100.0%; Pred. No. 7.1e-97;										
Matches 392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
QY	1	CGCGCGGAGCTGTGAGCGGGGACTCGGGTCCCTGAGGTCGTGATCTTCTCCGCTACT	60							
Db	665	CGCGCGGAGCTGTGAGCGGGGACTCGGGTCCCTGAGGTCGTGATCTTCTCCGCTACT	724							
QY	61	GAGACACGGCGGTAGGTCCACAGGCAGATCCAACCTGGGAGTTGAAGTGTGAGTGAGAGT	120							
Db	725	GAGACACGGCGGTAGGTCCACAGGCAGATCCAACCTGGGAGTTGAAGTGTGAGTGAGAGT	784							
QY	121	GAAGAGGAACACAGAGCTTCCGAGGGTTGTGTGTCAGTCACTCAGAGTCAGAAGGCC	180							
Db	785	GAAGAGGAACACAGAGCTTCCGAGGGTTGTGTGTCAGTCACTCAGAGTCAGAAGGCC	844							
QY	181	CTCGAAGTCGTCTCCCTCTCATCGGTGCCAGCCCATGGACCTTCTTCTCTCGTCACG	240							
Db	845	CTCGAAGTCGTCTCCCTCTCATCGGTGCCAGCCCATGGACCTTCTTCTCTCGTCACG	904							
QY	241	GCCATACTAGGAGGAAGAGGCCCGCAGGAGTGGAGGGGCTCAGCGAAGCTGGGGTGC	300							
Db	905	GCCATACTAGGAGGAAGAGGCCCGCAGGAGTGGAGGGGCTCAGCGAAGCTGGGGTGC	964							
QY	301	TGTTGGGGGTATCCGAGTCCAGAAACACCTGGAACCCGACAGAGATTTCTGGACTCCC	360							
Db	965	TGTTGGGGGTATCCGAGTCCAGAAACACCTGGAACCCGACAGAGATTTCTGGACTCCC	1024							
QY	361	CAGACGGGACCAGGAGGAGGACGGCATGACGC	392							
Db	1025	CAGACGGGACCAGGAGGAGGACGGCATGACGC	1056							
RESULT 5										
AC025553/c										
LOCUS	AC025553	180859 bp	DNA	linear HTG 01-SEP-2000						

DEFINITION	Homo sapiens chromosome X clone RP11-485B17, WORKING DRAFT
ACCESSION	SEQUENCE, 23 unordered pieces.
VERSION	AC025553
KEYWORDS	AC025553.5 GI:9958270
SOURCE	HTG; HTGS_PHASE1; HTGS_DRAFT.
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL	Waterston,R.H.
AUTHORS	The sequence of Homo sapiens clone
REFERENCE	Unpublished
TITLE	2 (bases 1 to 180859)
JOURNAL	Waterston,R.H.
AUTHORS	Direct Submission
TITLE	Submitted (10-MAR-2000) Genome Sequencing Center, Washington
JOURNAL	University School of Medicine, 4444 Forest Park Parkway, St. Louis,
COMMENT	MO 63108, USA
	On Sep 1, 2000 this sequence version replaced gi:9954853.
----- Genome Center -----	
Center: Washington University Genome Sequencing Center	
Center code: WUGSC	
Web site:http://genome.wustl.edu/gsc/index.shtml	
----- Project Information -----	
Center project name: H.NH0485B17	
----- Summary Statistics -----	
Sequencing vector: M13; 100%	
Sequencing vector: plasmid; 0%	
Chemistry: Dye-terminator Big Dye; 0% of reads	
Assembly program: Phrap; version 0.990319	
Consensus quality: 170599 bases at least Q40	
Consensus quality: 173308 bases at least Q30	
Consensus quality: 174949 bases at least Q20	
Insert size: 187000; agarose-fp	
Insert size: 178075; sum-of-contigs	
Quality coverage: 4.89 in Q20 bases; agarose-fp	
Quality coverage: 5.22 in Q20 bases; sum-of-contigs	

* NOTE: This is a 'working draft' sequence. It currently	
* consists of 23 contigs. The true order of the pieces	
* is not known and their order in this sequence record is	
* arbitrary. Gaps between the contigs are represented as	
* runs of N, but the exact sizes of the gaps are unknown.	
* This record will be updated with the finished sequence	
* as soon as it is available and the accession number will	
* be preserved.	
*	1 1341: contig of 1341 bp in length
*	1342 1441: gap of unknown length
*	1442 2824: contig of 1383 bp in length
*	2825 2924: gap of unknown length
*	2925 4091: contig of 1167 bp in length
*	4092 4191: gap of unknown length
*	4192 5300: contig of 1109 bp in length
*	5301 5400: gap of unknown length
*	5401 6531: contig of 1131 bp in length
*	6532 7946: contig of 1315 bp in length
*	7947 8046: gap of unknown length
*	8047 9816: contig of 1770 bp in length
*	9817 9917: gap of unknown length
*	9917 11066: contig of 1150 bp in length
*	11067 11166: gap of unknown length
*	11167 12295: contig of 1129 bp in length
*	12296 12395: gap of unknown length
*	12396 13579: contig of 1184 bp in length
*	13580 13679: gap of unknown length
*	13680 15201: contig of 1522 bp in length
*	15202 15301: gap of unknown length
*	15302 16390: contig of 1089 bp in length
*	16391 16490: gap of unknown length

```
* 16491 18667: contig of 2177 bp in length
* 18668 18767: gap of unknown length
* 18768 20869: contig of 2102 bp in length
* 20869 20969: gap of unknown length
* 20970 23568: contig of 2599 bp in length
* 23569 23669: gap of unknown length
* 23669 24252: contig of 584 bp in length
* 24252 24353: gap of unknown length
* 24353 28943: contig of 4591 bp in length
* 28943 37619: gap of unknown length
* 37619 37719: contig of 8576 bp in length
* 37719 49927: gap of unknown length
* 49927 50028: contig of 12208 bp in length
* 50028 69101: gap of unknown length
* 69101 69201: contig of 19073 bp in length
* 69201 90261: gap of unknown length
* 90261 90361: contig of 21060 bp in length
* 90361 141786: gap of unknown length
* 141786 141887: contig of 51426 bp in length
* 141887 180859: gap of unknown length
* 180859: contig of 38973 bp in length.
FEATURES
source
1..180859
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/clone="RP11-485B17"
1..1341
/note="assembly_name:Contig19"
1442..2824
/note="assembly_name:Contig24"
2925..4091
/note="assembly_name:Contig25"
4192..5300
/note="assembly_name:Contig26"
5401..6531
/note="assembly_name:Contig40"
6632..7946
/note="assembly_name:Contig44"
8047..9816
/note="assembly_name:Contig45"
9917..11066
/note="assembly_name:Contig46"
11167..12295
/note="assembly_name:Contig47"
12396..13579
/note="assembly_name:Contig48"
13680..15201
/note="assembly_name:Contig49"
15302..16390
/note="assembly_name:Contig50"
16491..18667
/note="assembly_name:Contig52"
18768..20869
/note="assembly_name:Contig53"
20970..23568
/note="assembly_name:Contig54
clone_end:SP6
vector_side:right"
23669..24252
/note="assembly_name:Contig6"
24353..28943
/note="assembly_name:Contig55"
29044..37619
/note="assembly_name:Contig56"
37720..49927
/note="assembly_name:Contig57"
50028..69100
/note="assembly_name:Contig58"
69201..90260
/note="assembly_name:Contig59"
90361..141786
/note="assembly_name:Contig60"
```

```
misc_feature 141887..180859
/note="assembly_name:Contig61"
BASE COUNT 53372 a 37170 c 37454 g 50639 t 2224 others
ORIGIN
Query Match 50.2%; Score 392; DB 2; Length 180859;
Best Local Similarity 100.0%; Pred. No. 8.6e-97;
Matches 392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCGGCGAGCTGTGAGCCGCGGACTCGGGTCCCTGAGGCTCTGATTCCTTCTCCGCTACT 60
Db 157680 GCGGCGAGCTGTGAGCCGCGGACTCGGGTCCCTGAGGCTCTGATTCCTTCTCCGCTACT 157621
QY 61 GACACACGGCGGTAGTCCACAGGCGAGATCCAACTGGGAGTTGAAGTGTGAGTGTGAGACT 120
Db 157620 GACACACGGCGGTAGTCCACAGGCGAGATCCAACTGGGAGTTGAAGTGTGAGTGTGAGACT 157561
QY 121 GAAGAGAACCCAGCAGGCTTCCGGAGGGTTGTGTGTGCTAGTGTGCTGAGTGTGAGAGGCC 180
Db 157560 GAAGAGAACCCAGCAGGCTTCCGGAGGGTTGTGTGTGCTAGTGTGCTGAGTGTGAGAGGCC 157501
QY 181 CTGGAAGTCGTCTGCTCTCATCGGTGCCAGGCCCATGGACCTTCTTGTCTCTGTCACG 240
Db 157500 CTGGAAGTCGTCTGCTCTCATCGGTGCCAGGCCCATGGACCTTCTTGTCTCTGTCACG 157441
QY 241 GCCATAACTAGGAGGAGGAGGCGCGAGAGTGGAGGGCTCAGGCGAAGCTGGGGTGC 300
Db 157440 GCCATAACTAGGAGGAGGAGGCGCGAGAGTGGAGGGCTCAGGCGAAGCTGGGGTGC 157381
QY 301 TGTGTGGGGTATCCGAGTCCCAAGACACCTGGAACCCCGACAGAGATTCTGGACTCCC 360
Db 157380 TGTGTGGGGTATCCGAGTCCCAAGACACCTGGAACCCCGACAGAGATTCTGGACTCCC 157321
QY 361 CAGACGGGACAGGAGGAGGCGCATGACGG 392
Db 157320 CAGACGGGACAGGAGGAGGCGCATGACGG 157289
RESULT 6
AX245237 LOCUS AX245237 396 bp DNA linear PAT 28-SEP-2001
DEFINITION Sequence 167 from Patent WO0166753.
ACCESSION AX245237
VERSION AX245237.1 GI:15859911
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 396)
AUTHORS Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D.,
Sudduth-Klinger,J., Reinhard,C., Randazzo,F., Kennedy,G.C., Pot,D.,
Kassam,A., Lamson,G., Drmanac,R., Crkvenjakov,R., Dickson,M.,
Drmanac,S., Iabat,I., Leshkowitz,D., Kita,D., Garcia,V. and
Stache-Grain,B.
TITLE Human genes and gene expression products
JOURNAL Patent: WO 0166753-A 167 13-SEP-2001;
FEATURES Chiron Corporation (US) ; Hyseq Inc. (US)
source Location/Qualifiers
1..396
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 82 a 97 c 146 g 71 t
ORIGIN
Query Match 50.1%; Score 391; DB 6; Length 396;
Best Local Similarity 100.0%; Pred. No. 1.1e-96;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 GCGGCGAGCTGTGAGCCGCGGACTCGGGTCCCTGAGGCTCTGATTCCTTCTCCGCTACTG 61
Db 1 CCGGCGAGCTGTGAGCCGCGGACTCGGGTCCCTGAGGCTCTGATTCCTTCTCCGCTACTG 60
```



```
QY 62 AGACGGCGGCTAGGTCCACAGGAGATCCAACTGGGAGTTGAAGTGTGAGTGAGAGTG 121
|||||
Db 61 AGACGGCGGCTAGGTCCACAGGAGATCCAACTGGGAGTTGAAGTGTGAGTGAGAGTG 120
|||||
QY 122 AGAGAAACACAGAGGCTTCGGAGGGTGTGTGCTGAGTCACTCAGAGTGAGAGGCC 181
|||||
Db 121 AAGAGAAACACAGAGGCTTCGGAGGGTGTGTGCTGAGTCACTCAGAGTGAGAGGCC 180
|||||
QY 182 TCGAAGTGTCTGCTCTCTCATGCGGTGCCACGCCATGGACCTTCTGCTCGTCACGG 241
|||||
Db 181 TCGAAGTGTCTGCTCTCTCATGCGGTGCCACGCCATGGACCTTCTGCTCGTCACGG 240
|||||
QY 242 CCATAACTAGGAGGAGAGGCCGAGGAGTGGAGGGCTCAGCGAAGCTGGGGTCT 301
|||||
Db 241 CCATAACTAGGAGGAGAGGCCGAGGAGTGGAGGGCTCAGCGAAGCTGGGGTCT 300
|||||
QY 302 GTTGGGGTATCCGAGTCCAGAACACCTGGAACCCCGACAGAGATCTTGAGATCCCC 361
|||||
Db 301 GTTGGGGTATCCGAGTCCAGAACACCTGGAACCCCGACAGAGATCTTGAGATCCCC 360
|||||
QY 362 AGACGGGACGAGGAGGAGCGGATGAGCG 392
|||||
Db 361 AGACGGGACGAGGAGGAGCGGATGAGCG 391
|||||

RESULT 7
AF251237
LOCUS AF251237 Homo sapiens XAGE-1 mRNA, complete cds. linear PRI 23-AUG-2000
DEFINITION
ACCESSION AF251237
VERSION AF251237.1 GI:9885324
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE
AUTHORS Liu,X.F., Helman,L.J., Yeung,C., Lee,B. and Pastan,I.
TITLE XAGE-1, A New Gene That is Frequently Expressed in Ewing's Sarcoma
JOURNAL Unpublished
AUTHORS Liu,X.F., Helman,L.J., Yeung,C., Lee,B. and Pastan,I.
TITLE Direct Submission
JOURNAL Submitted (31-MAR-2000) Lab of Mol. Biol., NCI, NIH, 37 Convent Dr.
MSC 4255, Bethesda, MD 20892, USA
FEATURES
source
location/Qualifiers
1..626
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/tissue_type="testis"
86..526
/note="expressed in Ewing's sarcoma"
/codon_start=1
/product="XAGE-1"
/protein_id="AAG01401.1"
/db_xref="GI:9885325"
/translation="MLLWCPQACSLGVFPSPVWGTTRRSCPEATRPVPEWILSP
LLRHGHGTQNTHTSPSPSPESPKKNQOLKVLHLSRQKRIQLRSQCATWK
VICKSISQTPGPNLDLGSQVKVLIIPKEEHRMPAGEEQPV"
BASE COUNT 194 a 153 c 170 g 109 t
ORIGIN

Query Match 50.1%; Score 391; DB 9; Length 626;
Best Local Similarity 98.7%; Pred. No. 1.2e-96;
Matches 394; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 383 GGCATGAGCGACACACAAACACAGAACCCAGGCGAGTCCAGGAGGCCAGTAATGG 442
|||||
Db 225 GACAGGGGACACACAAACACAGAACCCAGGCGAGTCCAGGAGGCCAGTAATGG 284
|||||
QY 443 AGAGCCCCAAAGAACAGCAGCTGAAAGTGGGATCCTACACCTGGGCGAGAG 502
|||||
```

```
Db 285 AGAGCCCCAAAGAACAGCAGCTGAAAGTCCGGATCCTACACCTGGCAGCAGAC 344
QY 503 AGAAGAAGATCAGGATACAGCTGAGATCCCAGTGGCGACATGGAAGTGATCTGCAAGA 562
|||||
Db 345 AGAAGAAGATCAGGATACAGCTGAGATCCCAGTGGCGACATGGAAGTGATCTGCAAGA 404
|||||
QY 563 GCTGCATCAGTCAAAACACCGGGGATAAAATCTGGATTTCCGGCTCAAGGTGAAGA 622
|||||
Db 405 GCTGCATCAGTCAAAACACCGGGGATAAAATCTGGATTTCCGGCTCAAGGTGAAGA 464
|||||
QY 623 TAATACCTAAAGAGAACACTGTAAATGCCAGAGCAGGTGAAGAGCAACCAAGTTT 682
|||||
Db 465 TAATACCTAAAGAGAACACTGTAAATGCCAGAGCAGGTGAAGAGCAACCAAGTTT 524
|||||
QY 683 AAATGAACACAGCTGAACCAACCGAGCTGGTTTATATATAGATATTTGACTTAAACTA 742
|||||
Db 525 AAATGAACACAGCTGAACCAACCGAGCTGGTTTATATATAGATATTTGACTTAAACTA 584
|||||
QY 743 TCTCAATAAAGTTTTCAGCTTTTCACCAAAAAA 781
|||||
Db 585 TCTCAATAAAGTTTTCAGCTTTTCACCAAAAAA 623
|||||

RESULT 8
AX062442
LOCUS AX062442 Sequence 69 from Patent WO0100828. linear PAT 24-JAN-2001
DEFINITION
ACCESSION AX062442
VERSION AX062442.1 GI:12540317
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE
AUTHORS Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S., Carter,D., Retter,M.W. and Mannion,J.
TITLE Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL Patent: WO 0100828-A 69 04-JAN-2001; CORIXA CORPORATION (US)
FEATURES
source
location/Qualifiers
1..399
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 150 a 87 c 94 g 67 t 1 others
ORIGIN

Query Match 50.0%; Score 390.6; DB 6; Length 399;
Best Local Similarity 98.5%; Pred. No. 1.5e-96;
Matches 393; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 383 GGCATGAGCGACACACAAACACAGAACCCAGGCGAGTCCCGAGGCGCCAGTAATGG 442
|||||
Db 1 GACACGGCGACACACAAACACAGAACCCAGGCGAGTCCCGAGGCGCCAGTAATGG 60
|||||
QY 443 AGAGCCCCAAAGAACAGCAGCTGAAAGTGGGATCCTACACCTGGGCGAGAG 502
|||||
Db 61 AGAGCCCCAAAGAACAGCAGCTGAAAGTGGGATCCTACACCTGGGCGAGAG 120
|||||
QY 503 AGAAGAAGATCAGGATACAGCTGAGATCCCGAGCATGGAGGTGATCTGCAAGA 562
|||||
Db 121 AGAAGAAGATCAGGATACAGCTGAGATCCCGAGCATGGAGGTGATCTGCAAGA 180
|||||
QY 563 GCTGCATCAGTCAAAACACCGGGGATAAAATCTGGATTTCCGGCTCAAGGTGAAGA 622
|||||
Db 181 GCTGCATCAGTCAAAACACCGGGGATAAAATCTGGATTTCCGGCTCAAGGTGAAGA 240
|||||
QY 623 TAATACCTAAAGAGAACACTGTAAATGCCAGAGCAGGTGAAGAGCAACCAAGTTT 682
|||||
Db 241 TAATACCTAAAGAGAACACTGTAAATGCCAGAGCAGGTGAAGAGCAACCAAGTTT 300
|||||
QY 683 AAATGAACAGCTGAACCAACCGAGCTGGTTTATATATAGATATTTGACTTAAACTA 742
|||||
```

||||| 301 AAATGAAGACAAGCTGAACAACCAAGCTGGTTTATATTAGATATTGACTTAAACTA 360

QY 743 TCTCAATAAAGTTTTGACAGCTTCCACCAAAAAA 781

Db 361 TCTCAATAAAGTTTTGACAGCTTCCACCAAAAAA 399

RESULT 9

HSA290447

LOCUS Homo sapiens mRNA for 9 kD cancer/testis associated protein 05-MAY-2001

DEFINITION (Xage-lb).

ACCESSION AJ290447

VERSION AJ290447.1 Gi:13992498

KEYWORDS cancer/testis associated protein; CTP9 gene.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 463)

Zendman,A.J.W., van Kraats,A.A., Weidle,U.H., Ruiter,D.J. and Van Muijen,G.N.P.

TITLE CTP9, a novel cancer/testis-associated gene, induced in melanoma metastases

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 463)

AUTHORS Zendman,A.J.W.

TITLE Direct Submission

JOURNAL Submitted (04-MAY-2000) Zendman A.J.W., Pathology, University Medical Center St. Radboud, Geert Grooteplein zuid 24, Nijmegen, 6500 HB POBox 9101, NETHERLANDS

REMARK Revised by author 22-JAN-2001

FEATURES

source

1..463

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="X"

/map="Xp11.21-Xp11.22"

/tissue_type="melanoma metastasis"

125..370

/gene="Xage-lb"

125..370

/gene="Xage-lb"

/codon_start=1

/product="9 kD cancer/testis associated protein"

/protein_id="CAC38107.1"

/db_xref="GI:13992499"

/translation="MESPKKKNOOLKVGLHLGSPKKIRIQLRSQCATWKVICKSCI

SOTPGINLDLGSYKVKLIIPKEECHKMPAGEEQPV"

BASE COUNT 155 a 106 c 120 g 82 t

ORIGIN

Query Match 49.6%; Score 387; DB 9; Length 463;

Best Local Similarity 98.7%; Pred. No. 1.5e-95;

Matches 390; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 383 GGCATGAGGCACACACAAAACACAGACCCACAGCCAGTCCCGAGGCCAGTAAATGG 442

Db 69 GACACGGCGGACACACAAAACACAGAACCCACAGCCAGTCCCGAGGCCAGTAAATGG 128

QY 443 AGAGCCCCAAAAGAACACAGACCTGAAAGTCGGGATCCCTACACCTGGGCAGAC 502

Db 129 AGAGCCCCAAAAGAACACAGACCTGAAAGTCGGGATCCCTACACCTGGGCAGAC 188

QY 503 AGAAGAGATCAGGATACAGCTGAGATCCAGTCGCGGACATGGAAGGTGATCTGCAAGA 562

Db 189 AGAAGAGATCAGGATACAGCTGAGATCCAGTCGCGGACATGGAAGGTGATCTGCAAGA 248

QY 563 GCTGCATCAGTCAAAACACCGGGATAAATCTGGATTGGGTTCCCGCGCTCAAGGTGAAGA 622

Db 249 GCTGCATCAGTCAAAACACCGGGATAAATCTGGATTGGGTTCCCGCGCTCAAGGTGAAGA 308

QY 623 TAATACCTAAAGAGGAACTGTAAATGCCAGAAGCAGGTGAAGAGCAACCAACTTT 682

Db 309 TAATACCTAAAGAGGAACTGTAAATGCCAGAAGCAGGTGAAGAGCAACCAACTTT 368

QY 683 AAATGAAGACAAGCTGAAACAACCAAGCTGGTTTTATATTAGATATTGACTTAAACTA 742

Db 369 AAATGAAGACAAGCTGAAACAACCAAGCTGGTTTTATATTAGATATTGACTTAAACTA 428

QY 743 TCTCAATAAAGTTTTGACAGCTTCCACCAAAAAA 777

Db 429 TCTCAATAAAGTTTTGACAGCTTCCACCAAAAAA 463

RESULT 10

AX063163

LOCUS Sequence 790 from Patent WO0100828.

DEFINITION AX063163

ACCESSION AX063163

VERSION AX063163.1 GI:12541017

KEYWORDS human.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 457)

Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S., Carter,D., Retter,M.W. and Mannion,J.

TITLE Compositions and methods for the therapy and diagnosis of lung cancer

JOURNAL Patent: WO 0100828-A 790 04-JAN-2001;

FEATURES

source

1..457

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 146 a 105 c 121 g 85 t

ORIGIN

Query Match 47.2%; Score 369; DB 6; Length 457;

Best Local Similarity 98.7%; Pred. No. 1.3e-90;

Matches 372; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 383 GGCATGAGGCACACAAAACACAGACCCACAGCCAGTCCCGAGGCCAGTAAATGG 442

Db 81 GACACGGCGGACACACAAAACACAGAACCCACAGCCAGTCCCGAGGCCAGTAAATGG 140

QY 443 AGAGCCCCAAAAGAACACAGCAGCTGAAAGTCGGGATCCCTACACCTGGGCAGAC 502

Db 141 AGAGCCCCAAAAGAACACAGCAGCTGAAAGTCGGGATCCCTACACCTGGGCAGAC 260

QY 503 AGAAGAGATCAGGATACAGCTGAGATCCAGTCGCGGACATGGAAGGTGATCTGCAAGA 56

Db 201 AGAAGAGATCAGGATACAGCTGAGATCCAGTCGCGGACATGGAAGGTGATCTGCAAGA 260

QY 563 GCTGCATCAGTCAAAACACCGGGATAAATCTGGATTGGGTTCCCGCGCTCAAGGTGAAGA 622

Db 261 GCTGCATCAGTCAAAACACCGGGATAAATCTGGATTGGGTTCCCGCGCTCAAGGTGAAGA 320

QY 623 TAATACCTAAAGAGGAACTGTAAATGCCAGAAGCAGGTGAAGAGCAACCAACTTT 682

Db 321 TAATACCTAAAGAGGAACTGTAAATGCCAGAAGCAGGTGAAGAGCAACCAACTTT 380

QY 683 AAATGAAGACAAGCTGAAACAACCAAGCTGGTTTTATATTAGATATTGACTTAAACTA 742

Db 381 AAATGAAGACAAGCTGAAACAACCAAGCTGGTTTTATATTAGATATTGACTTAAACTA 440

QY 743 TCTCAATAAAGTTTTGC 759

Db 441 TCTCAATAAAGTTTTGC 457

RESULT 11

HSA318879

LOCUS HSA318879 481 bp mRNA linear PRI 14-JAN-2002
DEFINITION Homo sapiens mRNA for XAGE-1d protein.
ACCESSION AJ318879
VERSION AJ318879.1 GI:18157207
KEYWORDS XAGE-1d gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Zandman, A.J.W., van Kraats, A.A., Weidle, U.H., Ruiter, D.R. and Van Muijen, G.N.P.
TITLE Expression profile of members of the XAGE cancer/testis antigen family
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 481)
AUTHORS Zandman, A.J.W.
TITLE Direct Submission
JOURNAL Submitted (14-AUG-2001) Zandman A.J.W., Department of Pathology, University Medical Center St Radboud, Geert Grooteplein Zuid 24, PO Box 9101 6500 HB Nijmegen, NETHERLANDS
FEATURES
source
1..481
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp11.21-Xp11.22"
/tissue_type="Ewing sarcoma"
125..334
/gene="XAGE-1d"
125..334
/gene="XAGE-1d"
/codon_start=1
/product="XAGE-1d protein"
/protein_id="CAC82987.1"
/db_xref="GI:18157208"
/translation="MESPKKKNOOLKVGILHLSRQKKIRIQLRSQVLGRNMRDWD
LQELHOSNTGDKSGFRRQEDNT"
BASE COUNT 163 a 107 c 127 g 84 t
ORIGIN
Query Match 46.5%; Score 363; DB 9; Length 481;
Best Local Similarity 94.9%; Pred. No. 5.9e-89;
Matches 392; Conservative 0; Mismatches 5; Indels 16; Gaps 1;
QY 383 GGCATGAGCGCACACACACACAGAACACACAGCCAGTCCAGGAGCCAGTAATGG 442
Db 69 GACACGGCGACACACACACACACAGAACACACAGCCAGTCCAGGAGCCAGTAATGG 128
QY 443 AGAGCCCCAAAAGAGAACACAGCTGAAAGTCGGGATCCTACACCTGGGCAGCAGAC 502
Db 129 AGAGCCCCAAAAGAGAACACAGCTGAAAGTCGGGATCCTACACCTGGGCAGCAGAC 188
QY 503 AGAAGAAGATCAGGATACAGCTGAGATCCAG-----TGC CGCAGCATGG 546
Db 189 AGAAGAAGATCAGGATACAGCTGAGATCCAGTCTGGGAAGGGAATGCGCGCATGG 248
QY 547 AAGGTGATCTGAAGAGCTGTCATAGTCAACACCGGGGATAAATCTGGATTGGGTTC 606
Db 249 AAGGTGATCTGAAGAGCTGTCATAGTCAACACCGGGGATAAATCTGGATTGGGTTC 308
QY 607 GCGTCAGGTCAAGATTAATACCTAAAGAGGACACTGTAAATCCAGAGCAGGTGAA 666
Db 309 GCGTCAGGTCAAGATTAATACCTAAAGAGGACACTGTAAATCCAGAGCAGGTGAA 368
QY 667 GAGCAACCAAGTTTAAATGAAGCAAGCTGAAACACGCAAGCTGTTTTATATTAGA 726
Db 369 GAGCAACCAAGTTTAAATGAAGCAAGCTGAAACACGCAAGCTGTTTTATATTAGA 428
QY 727 TATTTGACTTAACATCTCAATAAAGTTTTCAGCTTTCACCAAAAAAAA 779
Db 429 TATTTGACTTAACATCTCAATAAAGTTTTCAGCTTTCACCAAAAAAAA 481

RESULT 12
AX062443
LOCUS
DEFINITION Sequence 70 from Patent WO0100828.
ACCESSION AX062443
VERSION AX062443.1 GI:12540318
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 479)
AUTHORS Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedwick, T.S., Carter, D., Retter, M.W. and Mannion, J.
TITLE Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL Patent: WO 0100828-A 70 04-JAN-2001;
CORIXA CORPORATION (US)
FEATURES
source
1..479
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 163 a 107 c 125 g 84 t
ORIGIN
Query Match 45.3%; Score 354; DB 6; Length 479;
Best Local Similarity 94.7%; Pred. No. 1.8e-86;
Matches 394; Conservative 0; Mismatches 5; Indels 17; Gaps 2;
QY 383 GGCATGAGCGCACACACACACAGAACACACAGCCAGTCCAGGAGCCAGTAATGG 442
Db 64 GACACGGCGACACACACACACAGAACACACAGCCAGTCCAGGAGCCAGTAATGG 123
QY 443 AGAGCCCCAAAAGAGAACACAGCTGAAAGTCGGGATCCTACACCTGGGCAGCAGAC 502
Db 124 AGAGCCCCAAAAGAGAACACAGCTGAAAGTCGGGATCCTACACCTGGGCAGCAGAC 183
QY 503 AGAAGAAGATCAGGATACAGCTGAGATCCAG-----TGC CGCAGCATGG 546
Db 184 AGAAGAAGATCAGGATACAGCTGAGATCCAGTCTGGGAAGGGAATGCGCGCATGG 243
QY 547 AAGGTGATCTGAAGAGCTGTCATAGTCAACACCGGGGATAAATCTGGATTGGGTTC 606
Db 244 AAGGTGATCTGAAGAGCTGTCATAGTCAACACCGGGGATAAATCTGGATTGGGTTC 303
QY 607 GCGTCAGGTCAAGATTAATACCTAAAGAGGAACTGTAAATCCAGAGCAGGTGAA 666
Db 304 GCGTCAGGTCAAGATTAATACCTAAAGAGGAACTGTAAATCCAGAGCAGGTGAA 363
QY 667 GAGCAACCAAGTTTAAATGAAGCAAGCTGAAACACGCAAGCTGTTTTATATTAGA 725
Db 364 GAGCAACCAAGTTTAAATGAAGCAAGCTGAAACACGCAAGCTGTTTTATATTAGA 423
QY 726 ATATTTGACTTAACATCTCAATAAAGTTTTCAGCTTTCACCAAAAAAAA 781
Db 424 ATATTTGACTTAACATCTCAATAAAGTTTTCAGCTTTCACCAAAAAAAA 479
RESULT 13
AX063165
LOCUS
DEFINITION Sequence 792 from Patent WO0100828.
ACCESSION AX063165
VERSION AX063165.1 GI:12541018
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 461)
AUTHORS Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedwick, T.S., Carter, D., Retter, M.W. and Mannion, J.

TITLE Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL Patent: WO 0100828-A 792 04-JAN-2001;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1..461
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 150 a 104 c 123 g 84 t
ORIGIN
Query Match 44.8%; Score 350; DB 6; Length 461;
Best Local Similarity 94.8%; Pred. No. 2.2e-85;
Matches 379; Conservative 0; Mismatches 5; Indels 16; Gaps 1;
QY 383 GGCATGAGCGACACACACAAACACAGAACACACAGCCAGTCCCGAGGAGCCAGTAATGG 442
Db 62 GACAGGCGGACACACACAAACACACAGAACACACAGCCAGTCCCGAGGAGCCAGTAATGG 121
QY 443 AGAGCCCAAAAGAGAACACAGCAGCTGAAAGTCGGGATCCTACACCTGGCGCAGCAGAC 502
Db 122 AGAGCCCAAAAGAGAACACAGCAGCTGAAAGTCGGGATCCTACACCTGGCGCAGCAGAC 181
QY 503 AGAAGAAGATCAGGATACAGCTGAGATCCCGAG-----TGCAGCAGATGG 546
Db 182 AGAAGAAGATCAGGATACAGCTGAGATCCCGAGTGTGGGAAGGGAATGCGCGACATGG 241
QY 547 AAGGTGATCTGCAAGAGCTGCATCAGTCAACACCGGGGATTAATCTGATTTGGGTTCC 606
Db 242 AAGGTGATCTGCAAGAGCTGCATCAGTCAACACCGGGGATTAATCTGATTTGGGTTCC 301
QY 607 GGCCTCAAGGTGAAGATATACCTAAAGAGGAACACTGTAAATGCGAGAAGCAGGTGAA 666
Db 302 GGCCTCAAGGTGAAGATATACCTAAAGAGGAACACTGTAAATGCGAGAAGCAGGTGAA 361
QY 667 GAGCAACCAAGTTAAATGAAGACAGCTGAAACACGCGAAGCTGGTTTATATATAGA 726
Db 362 GAGCAACCAAGTTAAATGAAGACAGCTGAAACACGCGAAGCTGGTTTATATATAGA 421
QY 727 TATTGTACTTAACTATCTCAATAAAGTTTTCGAGCTTTC 766
Db 422 TATTGTACTTAACTATCTCAATAAAGTTTTCGAGCTTTC 461
RESULT 14
AX246260 391 bp DNA linear PAT 28-SEP-2001
LOCUS
DEFINITION Sequence 1190 from Patent WO0166753.
ACCESSION AX246260
VERSION AX246260.1 GI:15860934
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 391)
AUTHORS Williams,I.T., Escobedo,J., Innis,M.A., Garcia,P.D.,
Sudduth-Klinger,J., Reinhard,C., Randazzo,F., Kennedy,G.C., Pot,D.,
Kassam,A., Lamson,G., Drmanac,R., Crkvenjakov,R., Dickson,M.,
Drmanac,S., Labat,I., Leshkowitz,D., Kita,D., Garcia,V. and
Stache-Crain,B.
TITLE Human genes and gene expression products
JOURNAL Patent: WO 0166753-A 1190 13-SEP-2001;
Chiron Corporation (US) ; Hyseq Inc. (US)
FEATURES Location/Qualifiers
source 1..391
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 82 a 109 c 117 g 83 t
ORIGIN
Query Match 39.1%; Score 305; DB 6; Length 391;
Best Local Similarity 88.2%; Pred. No. 5.4e-73;

Matches 343; Conservative 0; Mismatches 45; Indels 1; Gaps 1;
QY 2 CGCGGAGAGCTGTAGCGCGGAGCTCGGGTCCCTCAGCTCGGATTTCTTCCGCTACTG 61
Db 1 CTGCGGAGCTGTGACCCGCGGAGCTTGGGGCCCTGATGCTGGATTTCTTCCGATACTG 60
QY 62 AGACACGCGGGTAGGTCCACAGGCA-GATCCAACTGGGAGTTGAAAGTGTGAGTGAGAGT 120
Db 61 AGACACGCGCGGTAGGTCCACAGGCACTATCCAACTGGAAGTTCAATTTGTGAGTGAGAGT 120
QY 121 GAAGAGGAACCAAGGAGCTTCCGAGGAGTTGTGTGTCAGTGACTCAGAGTGAGAAAGGCC 180
Db 121 GAACAGGAACCTTCCGGAGGTTGTGTGGCCAGTGACTCAAAGTGAGAAAGGCC 180
QY 181 CTCGAAGTCGTCGCTCCTCTCATCGGTGCGCACCGCATGAGACCTTCTTGTCTCGTCAACG 240
Db 181 CTCGAAGTCGTCCTACGTCTCATCGGGGCGCTCGCCCATGTCCTTCTGTCTCGCTCG 240
QY 241 GCCATAACTAGGAGTGAAGAGGCGCGAGAGTGGAGGGCTCAGGGGAAGCTGGGGTGC 300
Db 241 GTCATAACTAGGAGGAACAGAGGCGCGAGGAGTGTAAAGGCTCCTCAAGCTTGGGTGC 300
QY 301 TGTGGGGGTATCTGAGTCCAGAGACCTTGGAAACCCGACAGAGAGATTCTGGACTCCC 360
Db 301 TGTGGGGGTATCCGAATCCCACTAGCACCTTGGAAACCCGACTGAAGACTCTGCACCTCCC 360
QY 361 CAGCGGACCAAGGAGAGGCGGATCA 389
Db 361 CACAGGAACCAAGGAGAGGATACGCCATCA 389
RESULT 15
HSA318881 HSA318881 493 bp mRNA linear PRI 14-JAN-2002
LOCUS Homo sapiens mRNA for XAGE-3 protein.
DEFINITION
ACCESSION AJ318881
VERSION AJ318881.1 GI:18157211
KEYWORDS XAGE-3 gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Zendman,A.J.W., van Kraats,A.A., Weidle,U.H., Rulter,D.R. and Van
Muijen,G.N.P.
TITLE Expression profile of members of the XAGE cancer/testis antigen
family
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 493)
AUTHORS Zendman,A.J.W.
TITLE Direct Submission
JOURNAL Submitted (14-AUG-2001) Zendman A.J.W., Department of Pathology,
University Medical Center St Radboud, Geert Grooteplein Zuid 24, PO
Box 9101 6500 HB Nijmegen, NETHERLANDS
FEATURES Location/Qualifiers
source 1..493
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp11.21-Xp11.22"
/tissue.type="Ewing sarcoma"
61..396
/gene="XAGE-3"
61..396
/gene="XAGE-3"
/codon_start=1
/product="XAGE-3 protein"
/protein_id="CAC83008.1"
/db_xref="GI:18157212"
/translation="MIWRGRSTYPRRBRSSVPPPELIGPMLPEDEBPQEPPTESR
DPAGQREEDQGAETQVPLEADLQELSOSKTTGGCGNGPDDQKILPKSEQFKMP
EGGRQPOV"
BASE COUNT 167 a 98 c 125 g 103 t


```
* 20562 20661: gap of 100 bp
* 20662 21456: contig of 795 bp in length
* 21457 21556: gap of 100 bp
* 21557 22382: contig of 826 bp in length
* 22383 22482: gap of 100 bp
* 22483 23306: contig of 824 bp in length
* 23307 23406: gap of 100 bp
* 23407 24217: contig of 811 bp in length
* 24218 24317: gap of 100 bp
* 24318 25116: contig of 799 bp in length
* 25117 25216: gap of 100 bp
* 25217 25999: contig of 783 bp in length
* 26000 26099: gap of 100 bp
* 26100 26879: contig of 780 bp in length
* 26880 26979: gap of 100 bp
* 26980 27773: contig of 794 bp in length
* 27774 27873: gap of 100 bp
* 27874 28668: contig of 795 bp in length
* 28669 28768: gap of 100 bp
* 28769 29592: contig of 824 bp in length
* 29593 29692: gap of 100 bp
* 29693 30523: contig of 831 bp in length
* 30524 30623: gap of 100 bp
* 30624 31423: contig of 800 bp in length
* 31424 31523: gap of 100 bp
* 31524 32321: contig of 798 bp in length
* 32322 32421: gap of 100 bp
* 32422 33244: contig of 823 bp in length
* 33245 33344: gap of 100 bp
* 33345 34172: contig of 828 bp in length
* 34173 34272: gap of 100 bp
* 34273 35079: contig of 807 bp in length
* 35080 35179: gap of 100 bp
* 35180 35967: contig of 788 bp in length
* 35968 36067: gap of 100 bp
* 36068 36846: contig of 779 bp in length
* 36847 36946: gap of 100 bp
* 36947 37740: contig of 794 bp in length
* 37741 37840: gap of 100 bp
* 37841 38651: contig of 811 bp in length
* 38652 38751: gap of 100 bp
* 38752 39566: contig of 815 bp in length
* 39567 39666: gap of 100 bp
* 39667 40472: contig of 806 bp in length
* 40473 40572: gap of 100 bp
* 40573 41381: contig of 809 bp in length
* 41382 41481: gap of 100 bp
* 41482 42296: contig of 815 bp in length
* 42297 42396: gap of 100 bp
* 42397 43213: contig of 817 bp in length
* 43214 43313: gap of 100 bp
* 43314 44124: contig of 811 bp in length
* 44125 44224: gap of 100 bp
* 44225 45028: contig of 804 bp in length
* 45029 45128: gap of 100 bp
* 45129 45937: contig of 809 bp in length
* 45938 46037: gap of 100 bp
* 46038 46814: contig of 777 bp in length
* 46815 46914: gap of 100 bp
* 46915 47710: contig of 796 bp in length
* 47711 47810: gap of 100 bp
* 47811 48580: contig of 770 bp in length
* 48581 48680: gap of 100 bp
* 48681 49465: contig of 785 bp in length
* 49466 49565: gap of 100 bp
* 49566 50377: contig of 812 bp in length
* 50378 50477: gap of 100 bp
* 50478 51260: contig of 783 bp in length
* 51261 51360: gap of 100 bp
* 51361 52143: contig of 783 bp in length
* 52144 52243: gap of 100 bp
* 52244 53049: contig of 806 bp in length
* 53050 53149: gap of 100 bp

* 53150 53959: contig of 810 bp in length
* 53960 54059: gap of 100 bp
* 54060 54832: contig of 773 bp in length
* 54833 54932: gap of 100 bp
* 54933 55712: contig of 780 bp in length
* 55713 55812: gap of 100 bp
* 55813 56604: contig of 792 bp in length
* 56605 56704: gap of 100 bp
* 56705 57487: contig of 783 bp in length
* 57488 57587: gap of 100 bp
* 57588 58395: contig of 808 bp in length
* 58396 58495: gap of 100 bp
* 58496 59302: contig of 807 bp in length
* 59303 59402: gap of 100 bp
* 59403 60205: contig of 803 bp in length
* 60206 60305: gap of 100 bp
* 60306 61119: contig of 814 bp in length
* 61120 61219: gap of 100 bp
* 61220 62026: contig of 807 bp in length
* 62027 62126: gap of 100 bp
* 62127 62941: contig of 815 bp in length
* 62942 63041: gap of 100 bp
* 63042 63857: contig of 816 bp in length
* 63858 63957: gap of 100 bp
* 63958 64772: contig of 815 bp in length
* 64773 64872: gap of 100 bp
* 64873 65675: contig of 803 bp in length

Query Match 26.4%; Score 205.8; DB 2: Length 81953;
Best Local Similarity 77.8%; Pred. No. 1.5e-45;
Matches 281; Conservative 0; Mismatches 62; Indels 18; Gaps 2;

QY 6 GGAGCTGTGAGCGCGGACTCGGGTCCCTGAGGTCTGGATTCTTCGCGCTACTCAGAC 65
||||| ||| ||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||
Db 78751 GGA CTGTGATCTCTGACTTCGGTCCCTGAGGTCTGGATTCTTCGCGCTACTCAGAC 78810

QY 66 ACGCGGGTAGTGTCCACAGGAGATCCAACTGGAGTTGAAGTGTGAGTGAGAGTGAAGA 125
||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||
Db 78811 GCAGCCAGTAGTGTCCACAGCGCATCAACTGGAGTTGAAGTGTGAGTGAGAGTGC 78870

QY 126 GGAACACAGCAGGCTCCGGAGGT-----TGTGTGTCTGACTGACTCAGAGTGAAGGCC 180
||||| ||||||| ||||||||| ||| ||| ||||||| ||||||||| |||||||||
Db 78871 GGAGCCAGCGGGCTTCGGAGGTTCCGGCGGCTGATTGTTGACTCCGAGTGAAGGCC 78930

QY 181 CTCGAAGTCTGCTCCCTCTCTCATGCGTGCACGCCCATGGACCTCTTCTCTCGTCACG 240
||| ||||||||| ||| ||||| ||| ||||| ||| ||||| ||||| ||| |||
Db 78931 CTTGATGTCGCTCCTCTCTGTCGCGCCCGCAGCATATGGGCTCTTCTCTTGTAGGG 78990

QY 241 GCCATTAAGGAGGAGGAGGCGCGAGGAGT-----GGAGGGCTCAGGC 287
||| ||| ||| ||||||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 78991 GTCAGGACGAAAGAGAGGAGGCTTCGGAGTGCACGCGNGGTTAGGTGAAGCTGGGC 79050

QY 288 GAAGCTGGGGTCTGTTGGGGTATCCGAGTCCGACAGCACCTGGAACCCGACAGAG 347
||||| ||||||||| ||||| ||||||||| ||||||||| ||||||||| |||||||||
Db 79051 AATGCTGGGGTCTGTTGGTGTATTCCAGTCCGACAGAGGCTGGAAGTCCGACAGAGG 79110

QY 348 A 348
+
Db 79111 A 79111

RESULT 19
AL445227
LOCUS
DEFINITION Homo sapiens chromosome X clone RP13-97115, *** SEQUENCING IN
PROGRESS ***, 6 unordered pieces.
ACCESSION AL445227
VERSION AL445227.7 GI:11611395
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```


<http://www.chori.org/bacpac/home.htm>

VECTOR: pBACe3.6

This sequence is the entire insert of clone RP13-77011 The true left end of clone RP13-97115 is at 15298 in this sequence. The true right end of clone RP13-97115 is at 13113 in this sequence.

FEATURES

source	location/qualifiers	repeat_region
	1. .149749	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/chromosome="X"	
	/map="p11.21-11.3"	
	/clone="RP13-77011"	
	/clone_lib="RPC1-13.1"	
	18. .1475	
repeat_region	/note="L1ME1 repeat: matches 3178. .4658 of consensus"	/note="L2 repeat: matches 2560. .2599 of consensus"
repeat_region	1581. .1842	18264. .18349
repeat_region	/note="L1ME1 repeat: matches 4789. .5069 of consensus"	/note="43 copies 2 mer ga 62% conserved"
repeat_region	1933. .2226	19489. .19797
repeat_region	/note="AluSx repeat: matches 1. .294 of consensus"	/note="AluJb repeat: matches 1. .306 of consensus"
repeat_region	2438. .3266	19826. .20044
repeat_region	3553. .3584	/note="L1M3 repeat: matches 5220. .5419 of consensus"
repeat_region	/note="16 copies 2 mer ac 84% conserved"	20045. .20175
repeat_region	3835. .4145	/note="L1M3 repeat: matches 5419. .5496 of consensus"
repeat_region	/note="AluJb repeat: matches 1. .310 of consensus"	20270. .20929
repeat_region	4173. .4486	/note="L1MA5A repeat: matches 5447. .6104 of consensus"
repeat_region	/note="AluSx repeat: matches 1. .312 of consensus"	20930. .21226
repeat_region	4982. .5096	/note="AluSx repeat: matches 1. .297 of consensus"
repeat_region	/note="MLT2E repeat: matches 1. .119 of consensus"	21227. .21417
repeat_region	5098. .5254	/note="L1MA5A repeat: matches 6104. .6294 of consensus"
repeat_region	/note="AluSx repeat: matches 2. .162 of consensus"	21418. .21461
repeat_region	5288. .5529	/note="L1M4 repeat: matches 4295. .4338 of consensus"
repeat_region	/note="L2 repeat: matches 2429. .2705 of consensus"	21500. .23928
repeat_region	5597. .5710	/note="L1PA5 repeat: matches 3696. .6143 of consensus"
repeat_region	/note="MIR repeat: matches 32. .151 of consensus"	24145. .24341
repeat_region	6178. .6471	/note="L1MEC repeat: matches 1820. .2012 of consensus"
repeat_region	/note="AluSg repeat: matches 1. .296 of consensus"	24385. .24681
repeat_region	7218. .7504	/note="AluSx repeat: matches 1. .296 of consensus"
repeat_region	/note="AluJb repeat: matches 2. .302 of consensus"	24718. .28764
repeat_region	7598. .7692	/note="L1PA4 repeat: matches 2108. .6144 of consensus"
misc_feature	/note="MIR repeat: matches 67. .169 of consensus"	28765. .28886
	8791. .9673	/note="FLAM_C repeat: matches 3. .124 of consensus"
	/note="CpG island"	28894. .28928
repeat_region	9015. .9136	/note="L1PA4 repeat: matches 2064. .2097 of consensus"
repeat_region	/evidence=not_experimental	28929. .29221
repeat_region	/note="61 copies 2 mer ct 69% conserved"	/note="AluSc repeat: matches 1. .296 of consensus"
repeat_region	9677. .9718	29222. .29377
repeat_region	/note="21 copies 2 mer tt 85% conserved"	29428. .29551
repeat_region	9745. .10036	/note="L2 repeat: matches 2621. .2749 of consensus"
repeat_region	/note="MER92C repeat: matches 272. .552 of consensus"	29553. .29644
repeat_region	10551. .10751	/note="AluJb repeat: matches 205. .296 of consensus"
repeat_region	/note="MER92B repeat: matches 2. .212 of consensus"	29648. .30097
repeat_region	10857. .10969	/note="MER87 repeat: matches 1. .469 of consensus"
repeat_region	/note="MIR repeat: matches 40. .154 of consensus"	30104. .30153
repeat_region	11521. .11613	/note="25 copies 2 mer ca 76% conserved"
repeat_region	/note="MIR repeat: matches 63. .156 of consensus"	31811. .31976
repeat_region	11629. .12187	/note="MIR repeat: matches 46. .204 of consensus"
repeat_region	/note="L2 repeat: matches 1779. .2341 of consensus"	32005. .32260
repeat_region	12438. .12548	/note="128 copies 2 mer ta 80% conserved"
repeat_region	/note="L1MA9 repeat: matches 6163. .6270 of consensus"	32903. .33194
repeat_region	12975. .13093	/note="AluJo repeat: matches 1. .293 of consensus"
repeat_region	/note="MIR repeat: matches 98. .247 of consensus"	33677. .34221
repeat_region	14083. .14381	/note="MER68A repeat: matches 1. .568 of consensus"
repeat_region	/note="AluSx repeat: matches 1. .293 of consensus"	34372. .35093
repeat_region	14906. .15126	/note="TIGGER1 repeat: matches 1650. .2308 of consensus"
repeat_region	/note="L2 repeat: matches 2364. .2569 of consensus"	35094. .35333
repeat_region	15142. .15170	/note="L1PB1 repeat: matches 5909. .6155 of consensus"
repeat_region	/note="MER84B repeat: matches 1. .215 of consensus"	35334. .35437
repeat_region	15366. .16028	/note="TIGGER1 repeat: matches 2308. .2418 of consensus"
repeat_region	/note="MER4D repeat: matches 359. .1017 of consensus"	35477. .35878
repeat_region	16097. .16211	/note="MLT1-INTERNAL repeat: matches 923. .1579 of consensus"
repeat_region	/note="L2 repeat: matches 2631. .2749 of consensus"	36082. .36740
repeat_region	16199. .16240	/note="MLT1-INTERNAL repeat: matches 201. .885 of consensus"
		36783. .37121
		/note="L2 repeat: matches 824. .1219 of consensus"
		37158. .37328
		/note="L1M4 repeat: matches 4640. .4816 of consensus"
		39953. .39978
		/note="13 copies 2 mer ca 100% conserved"
		40525. .40566
		/note="MIR repeat: matches 110. .151 of consensus"
		44063. .44377
		/note="AluJo repeat: matches 1. .299 of consensus"


```
/note="match: GSS: Em:B52857"
complement(29901..30381)
/note="match: GSS: Em:AQ039266"
29989..30082
/note="HERVL40 repeat: matches 5295..5367 of consensus"
30081..30112
/note="16 copies 2 mer tg 100% conserved"
30144..30513
/note="MT2FB repeat: matches 1..366 of consensus"
complement(31753..32234)
/note="match: GSS: Em:AQ209396"
31770..31935
/note="MIR repeat: matches 29..200 of consensus"
32266..32756
/note="match: GSS: Em:B47902"
32364..32727
/note="match: GSS: Em:AQ107228"
32412..32777
/note="LIPAB repeat: matches 5793..6159 of consensus"
32835..33112
/note="L2 repeat: matches 2255..2522 of consensus"
33291..33584
/note="MER39 repeat: matches 98..395 of consensus"
34330..34453
/note="FLAM_C repeat: matches 1..124 of consensus"
34597..34970
/note="LTR3 repeat: matches 1..413 of consensus"
complement(join(35694..36254,36581..36721))
/note="match: GSS: Em:AQ389427"
36267..36573
/note="Alusc repeat: matches 1..309 of consensus"
complement(36574..36730)
/note="match: GSS: Em:AQ181456"
38336..38377
/note="MIR repeat: matches 110..151 of consensus"
complement(38817..39110)
/note="match: STS: Em:HS220XG1"
41596..41706
/note="L1 repeat: matches 4663..4779 of consensus"
41765..41821
/note="L2 repeat: matches 1163..1219 of consensus"
41849..42117
/note="L2 repeat: matches 811..1102 of consensus"
42146..42812
/note="MLT1-INTERNAL repeat: matches 201..885 of
consensus"
43013..43410
/note="MLT1-INTERNAL repeat: matches 920..1322 of
consensus"
43499..43559
/note="TIGGER1 repeat: matches 2288..2359 of consensus"
43560..43799
/note="L1PB1 repeat: matches 5909..6155 of consensus"
43800..44089
/note="TIGGER1 repeat: matches 2078..2288 of consensus"
44132..44550
/note="TIGGER1 repeat: matches 1650..2094 of consensus"
44701..45249
/note="MER68A repeat: matches 1..568 of consensus"
45741..46034
/note="AluJo repeat: matches 1..293 of consensus"
46311..46342
/note="16 copies 2 mer ct 90% conserved"
```

```
Query Match      22.3%; Score 174; DB 9; Length 119024;
Best Local Similarity 72.7%; Pred. No. 8.6e-37;
Matches 272; Conservative 0; Mismatches 85; Indels 17; Gaps 3;
```

```
QY 14 GAGCCGGGACTCGGGTCCCTGAGTCTGGATCTTTCTCCGCTACTGAGACACGGCGG 73
|||||
DB 18724 GAGCTGTGAGGTCAAGTCTCGAGTCTGATCTTCTCTCTCTACTGAGAGCGAGG 18665
|||||
```

```
QY 74 TAGGTCACAGCAGACCACTGGGAGTTGAAGTGTGAGTCAGAGTAAGAGGAAACCAG 133
```

```
Db 18664 TAGGTCACAGGCCAATCCAACTGGGAGTTGAAGTGTGAGTGGGTTGAGGAGGAGCCAG 18605
QY 134 CAGGCTTCGGAGGGTTGT---GTGGTCAGTCACTCAGAGTGAGAAAGGCCCTCGAAGTCG 190
|||||
Db 18604 CAGGCTTCAGAGGGTTCATCAGGGGGGAGGAGACTCAGAGGGAGAGGGCCTGAGGTCT 18545
QY 191 TCGTCTCTCTCATCGGTGCCAGCCCATGGACCTTCTTGTCTCTGTCACGGCCATACCTA 250
|||||
Db 18544 TCATCTCTTTTCACATGCCGCCACAGCCAGGGGCTTCTT---TTTGTGTCAGGCCACA 18486
QY 251 GGAGGAAGGAGGGCCGAGGAGTGGAGGGCTCAGCGGAAGCTG-----GGG 297
|||||
Db 18485 GGAGGAAGGAGGGCCCTTGGCCAAAGGGGATCGGTTAAATGGGCACGCTGTTGTGGG 18426
QY 298 TCGTGTGGGGTATCCGAGTCCCAAGACACCTGGAAGCCGCCACAGAGATTTCTGGACT 357
|||||
Db 18425 TCGTGTGGAGGTATCTGAGTCCCAAGAAATGCCTCAACACATCAGAGATATCCAGTGA 18366
QY 358 CCCCACACGGGACC 371
Db 18365 CTCTGCCAGGACC 18352

RESULT 25
AX078298
LOCUS AX078298 580 bp DNA linear PAT 22-FEB-2001
DEFINITION Sequence 102 from Patent W00107471.
ACCESSION AX078298
VERSION AX078298.1 GI:13157989
KEYWORDS SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 580)
AUTHORS Hillman,J.L., Lal,P., Tang,Y.T., Yue,H., Au-Young,J., Bandman,O.,
Azizmai,Y., Yang,J., Lu,D.A., Baughn,M.R., Patterson,C. and Shah,P.
TITLE Cell cycle and proliferation proteins
JOURNAL Pat'ent: WO 0107471-A 102 01-FEB-2001;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
source 1..580
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 4764233CB1"
BASE COUNT 167 a 135 c 160 g 118 t
ORIGIN
```

```
Query Match      21.6%; Score 169; DB 6; Length 580;
Best Local Similarity 78.0%; Pred. No. 1.5e-35;
Matches 216; Conservative 0; Mismatches 35; Indels 1; Gap: 1;
```

```
QY 429 GAGCCCACTAATGGAGGCCCAAGAAAGAACAGCAGCTGAAAGTCGGGATCCTACA 488
|||||
```

```
Db 305 GAACCCACTGTAGAGAGCCTTAAGAGAGAACCACCCACTAAAGTCGGAACTCCTACA 364
|||||
```

```
QY 489 CTTGGGAGCAGACAGAGAAGATCAGGATCAGATCCCATGCCAGTCGGCGACATGCAA 548
|||||
```

```
Db 365 CTTGATCAGAGAGAGAAGATGATCAGGTGTCAGCTGAGATTCAAGTCTCACCTGGAA 424
|||||
```

```
QY 549 GTGTATCTGCAAGAGCTGCATCAGTCAACACCCGGGATAAATCTGGATTGGGTTCCGG 608
|||||
```

```
Db 425 GCCGATCTCCAGGAGCTATGTGACACAAGACTGGGGATGGATGTGAAGGT---GGTACTGA 483
|||||
```

```
QY 609 CTTCAAGTGAAGATAATACCTAAAGAGAACTGTAAAATGCCAGAGCAGGTGAAGA 668
|||||
```

```
Db 484 TGTCAAGGGAAGATTCTACCAAAAAGCAGACACTTTAAAATGCCAGAAGCAGGTGAAG 543
|||||
```

```
QY 669 GCACCAAGTTTAAATGAAGACAGCTGAACAAC 705
|||||
```

```
Db 544 GAAATCACAGGTTTAAAGGAAGATAAGCTGAACAAC 580
|||||
```

```
RESULT 26
AC002415/c
LOCUS
DEFINITION Homo sapiens chromosome X clone bWXD142, *** SEQUENCING IN PROGRESS
***, 6 unordered pieces.
ACCESSION AC002415
VERSION AC002415.1 GI:2323256
KEYWORDS HTG; HTGS_PHASE1.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 93419)
AUTHORS Chen,E., Brownstein,B.H., States,D.J., Schlessinger,D. and
Mazzarella,R.
TITLE Direct Submission
JOURNAL Unpublished (1997)
REFERENCE 2 (bases 1 to 93419)
AUTHORS Brownstein,B.H., States,D.J. and Mazzarella,R.
TITLE Direct Submission
JOURNAL Submitted (12-AUG-1997) Center for Genetics in Medicine, Box 8232,
Washington University School of Medicine, 4566 Scott Avenue, St.
Louis, MO 63110, USA
COMMENT Current status of this project is available at:
'http://genome.wustl.edu/cgm/seq_projects.html'
Submitted by:
Elison Chen,
Advanced Center for Genetic Technology,
Applied Biosystems Division of Perlin Elmer Corp.,
850 Lincoln Center Drive,
Foster City, CA 94404 USA
e-mail: elison@genseq.apldbio.com

and

Buddy Brownstein,
Center for Genetics in Medicine,
Washington University School of Medicine, Box 8232
4566 Scott Avenue,
St. Louis, MO 63110, USA
e-mail: buddy@genetics.wustl.edu

and

David J. States,
Institute for Biomedical Computing
Washington University in St. Louis
700 South Euclid Ave.
St. Louis, MO 63108 USA
e-mail: states@ibc.wustl.edu.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
* 1 9560: contig of 9560 bp in length
* 9561 25127: contig of 15567 bp in length
* 25128 45675: contig of 20548 bp in length
* 45676 73399: contig of 27724 bp in length
* 73400 82075: contig of 8676 bp in length
* 82076 93419: contig of 11344 bp in length.
FEATURES
source
Location/Qualifiers
1. .93419
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/clone="bWXD142"
BASE COUNT 25991 a 21189 c 20209 g 26030 t
ORIGIN
```

```
Query Match 21.5%; Score 167.8; DB 2; Length 93419;
Best Local Similarity 70.8%; Pred. No. 4.3e-35;
Matches 288; Conservative 0; Mismatches 97; Indels 22; Gaps 4;

Qy 10 CTGTGAGCCGCGCACTCGGCTCCTGAGGCTCGATTCTTTCTCCGCTACTGAGACACGG 69
|| |||| | | || || || || || || || || || || || || || || || || ||
Db 1923 CTTGGAGCTGTGAGGCCAAGTTCCTAAGGCTAGATTCTTTCTCTCTACTGATGACG 1864

Qy 70 CGGTAGGTCCACAGGCAGATCCAACTCGGAGTTCAAGTCTGAGTGTGAGTGTGAGAGGAA 129
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1863 CAGGTAGTCCACAGGCCAACTCCAACTGGGAGTTGAAGTGTGAGTGTGAGGTTGAGGAG 1804

Qy 130 CCAGCAGGCTTCCGAGGGT---TGCTGTGTGCTCAGTCTCAGAGTGTGAGAGGCCCTCGAA 186
||| |||| | | || || || || || || || || || || || || || || || || ||
Db 1803 CCAGTGGCTTCTGGAGGGTCACCGTGGGAGGAGACTCAGAGGGAGAGAGGGCTCGAG 1744

Qy 187 GTCTGTCGCTCCTCTCATCGGCTGCCACCCCATGAGACTTCTTGTCTCTGTCACGGCCATA 246
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1743 GTCTTCATCTTTATATGTGCACCTGCAGCCATGGGCTTCTT-TTTTGTGCGGGCCACA 1685

Qy 247 ACTAGGGAGGAGGCGCGGAGTGTGGGGCTCAGGCGCAAGCTGGGT----- 298
||| |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1684 ACTGGGAGGAGGAGGCGCTCGGCATGGGGGAGCTGGGTATGATGGGTGTCATCTG 1625

Qy 299 -----GCTGTTGGGGTATCCGAGTCCAGAGCACTTGGAAACCC-----GACAGAGA 348
|| || || || || || || || || || || || || || || || || || || || || ||
Db 1624 GGGGAGCTGTTGGAGGTATCTGAGTCCAGAAATGCCTCAAGCCCCCAAAAGAGGACAGA 1565

Qy 349 TTCTGGACTCCCGACAGCGGACCAGGAGGACGGCATGAGCGACA 395
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1564 TTGTAGACTCCCGAGGGGACCCAGGACGGGACGGGTGGGTGGCA 1518

RESULT 27
AL445227/c
LOCUS AL445227 114344 bp DNA linear HTG 10-JUL-2001
DEFINITION Homo sapiens chromosome X clone RP13-97I15, *** SEQUENCING IN
PROGRESS ***, 6 unordered pieces.
ACCESSION AL445227
VERSION AL445227.7 GI:11611395
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Melazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 114344)
AUTHORS McIay,K.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Dec 9, 2000 this sequence version replaced gi:11493333.
COMMENT ----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: b897115
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; 108752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 112482 bases at least Q40
Consensus quality: 113245 bases at least Q30
Consensus quality: 113597 bases at least Q20
Insert size: 113844; sum-of-contigs
Insert size: 122396; 9.8% error; agarose-fp
Quality coverage: 5.91x in Q20 bases; sum-of-contigs Quality
coverage: 5.72x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
```

* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved

* 1 13809: contig of 13809 bp in length
* 13810 13909: gap of 100 bp
* 13910 51507: contig of 37598 bp in length
* 51508 51607: gap of 100 bp
* 51608 65679: contig of 14072 bp in length
* 65680 65779: gap of 100 bp
* 65780 69064: contig of 3285 bp in length
* 69065 69184: gap of 100 bp
* 69185 101518: contig of 32354 bp in length
* 101519 101618: gap of 100 bp
* 101619 114344: contig of 12726 bp in length.

FEATURES		Location/Qualifiers	
source	1..114344	/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
misc_feature	1..13809	/clone_lib="RPCI-13.1"	
		/note="assembly_fragment:00241"	
misc_feature	fragment_chain:1	vector_side:left	
	clone_end:SP6	13910..51507	
misc_feature	1	/note="assembly_fragment:00680"	
		fragment_chain:1	
misc_feature	51608..65679	/note="assembly_fragment:00207"	
		fragment_chain:1	
misc_feature	65780..69064	/note="assembly_fragment:00300"	
		69165..101518	
misc_feature	101619..114344	/note="assembly_fragment:01024"	
		/note="assembly_fragment:01825"	
misc_feature	clone_end:T7	vector_side:right	
BASE COUNT	32658 a 24338 c 24232 g 32616 t	500 others	
ORIGIN			
Query Match	21.5%	Score 167.8; DB 2;	Length 114344;
Best Local Similarity	70.8%	Pred. No. 4.4e-35;	
Matches 288; Conservative	0; Mismatches 97; Indels 22; Gaps 4;		
QY	10	CTGTGAGCCGCGACTCGGGTCCCTGAGGCTGAGTTCCTTCGCTACTGACACGG	69
Db	54813	CTTGAGCTGTGAGGCCAAGTCTCTAAGGCTAGATTCTTCTCTCTACTGATGCAG	54754
QY	70	CGGTAGTCTCACAGGAGAGTCCAACTGGGAGTTGAAGTGTGAGTGTGAGTGAAGAGAA	129
Db	54753	CAGGTAGTCTCACAGGAGCAATCCAACTGGGAGTTGAAGTGTGAGTGTGAGGAGGAGG	54694
QY	130	CCAGCAGCTTCCGGAGGGT---TGCTGTGCTAGTCACTCAGAGTGTGAGAGGCGCTCGAA	186
Db	54693	CCAGTGGGCTTCTGGAGGGTCAACCGTGGGAGGAGACTCAGAGGAGAGGGCCCTCGAG	54634
QY	187	GTCTGCTGCTCTCATCGGTGCCAGCCCATGGACCTTCTGTCTGTCAGGGCCATA	246
Db	54633	GTCTTCATCTTTATATGTGGCACTGCAGCCATGGGCGTCTT-TTTTGTGGGGCCACA	54575
QY	247	ACTAGGAGGAGGAGCGCGAGGTGGAGGGGCTCAGCGAGCTGGGGT-----	298
Db	54574	ACTGGGAGGAGGAGGCGCTCGSCATGGGGGAGCTGGGTTATGATGGGTCCATACTG	54515
QY	299	-----GCTGTGGGGGTATCCGAGTCCCGAGACACCTGGAAACCC-----GACAGAAGA	348

Db	54514	GGGGAGCTGTTGGAGGTATCTAGTCCCAGAAATGCTCAGCCCCCAAGAGGACAGA	54455
QY	349	TTCTGGACTCCCCAGACGAGGACGAGGAGGAGCGCATGAGCGACA	395
Db	54454	TTGTAGACTCCCCAGGGGGACCCAGGACGCGTGGTGGCA	54408
RESULT 28			
AL445236	Human DNA sequence from clone RP13-77011 on chromosome		
LOCUS	Xp11.21-11.3, complete sequence.		
DEFINITION	AL445236		
ACCESSION	AL445236		
VERSION	AL445236.22 GI:13184376		
KEYWORDS	HTG.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 149749)		
	Howden,P.		
TITLE	Direct Submission		
	Submitted (08-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,		
JOURNAL	CB10 15A, UK. E-mail enquiries: humquery@sanger.ac.uk		
	requests: clonerequest@sanger.ac.uk		
COMMENT	On Mar 2, 2001 this sequence version replaced gi:13160313.		
	During sequence assembly data is compared from overlapping clones.		
	Where differences are found these are annotated as variations		
	together with a note of the overlapping clone name. Note that the		
	variation annotation may not be found in the sequence submission		
	corresponding to the overlapping clone, as we submit sequences with		
	only a small overlap as described above.		
	This sequence was finished as follows unless otherwise noted: all		
	regions were either double-stranded or sequenced with an alternate		
	chemistry or covered by high quality data (i.e., phred quality >=		
	30); an attempt was made to resolve all sequencing problems, such		
	as compressions and repeats; all regions were covered by at least		
	one plasmid subclone or more than one M13 subclone; and the		
	assembly was confirmed by restriction digest. The following		
	abbreviations are used to associate primary accession numbers given		
	in the feature table with their source databases: Em: EMBL; SW:		
	SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP		
	database can be found at		
	http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence		
	was generated from part of bacterial clone contigs of human		
	chromosome X, constructed by the Sanger Centre Chromosome X Mapping		
	Group. Further information can be found at		
	http://www.sanger.ac.uk/HGP/ChrX		
	RP13-77011 is from the library RP13-13.1 constructed by the group		
	of Pletier de Jong. For further details see		
	http://www.chori.org/bacpac/home.htm		
	VECTOR: pBACE3.6		
	This sequence is the entire insert of clone RP13-77011 The true		
	left end of clone RP13-97115 is at 15298 in this sequence. The true		
	right end of clone RP13-97115 is at 131153 in this sequence.		
FEATURES	Location/Qualifiers		
	1..149749		
source	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="X"		
	/map="p11.21-11.3"		
	/clone="RP13-77011"		
	/clone_lib="RPCI-13.1"		
repeat_region	18..1475		
	/note="L1ME1 repeat: matches 3178. .4658 of consensus"		
repeat_region	1581..1842		
	/note="L1ME1 repeat: matches 4789. .5069 of consensus"		
repeat_region	1933..2226		
	/note="AluX repeat: matches 1. .294 of consensus"		
repeat_region	2458..3266		
	/note="L1ME1 repeat: matches 5254. .6148 of consensus"		
repeat_region	3553..3584		
	/note="16 copies 2 mer ac 84% conserved"		

repeat_region	3835..4145 /note="AluJb repeat: matches 1. .310 of consensus"	repeat_region	24385..24681 /note="AluSx repeat: matches 1. .296 of consensus"
repeat_region	4173..4486 /note="AluSx repeat: matches 1. .312 of consensus"	repeat_region	24718..28764 /note="L1PA4 repeat: matches 2108. .6144 of consensus"
repeat_region	4982..5096 /note="MLR2E repeat: matches 1. .119 of consensus"	repeat_region	28765..28886 /note="FLAM_C repeat: matches 3. .124 of consensus"
repeat_region	5098..5254 /note="AluSq repeat: matches 2. .162 of consensus"	repeat_region	28894..28928 /note="L1PA4 repeat: matches 2064. .2097 of consensus"
repeat_region	5288..5529 /note="L2 repeat: matches 2429. .2705 of consensus"	repeat_region	28929..29221 /note="AluSc repeat: matches 1. .296 of consensus"
repeat_region	5597..5710 /note="MIR repeat: matches 32. .151 of consensus"	repeat_region	29222..29377 /note="L1MEC repeat: matches 2097. .2258 of consensus"
repeat_region	6178..6471 /note="AluSg repeat: matches 1. .296 of consensus"	repeat_region	29428..29551 /note="L2 repeat: matches 2621. .2749 of consensus"
repeat_region	7218..7504 /note="AluJo repeat: matches 2. .302 of consensus"	repeat_region	29553..29644 /note="AluJb repeat: matches 205. .296 of consensus"
repeat_region	7598..7692 /note="MIR repeat: matches 67. .169 of consensus"	repeat_region	29648..30097 /note="MER87 repeat: matches 1. .469 of consensus"
misc_feature	8791..9673 /note="CpG island" /evidence=not_experimental	repeat_region	30104..30153 /note="25 copies 2 mer ca 76% conserved"
repeat_region	9015..9136 /note="61 copies 2 mer ct 69% conserved"	repeat_region	31811..31976 /note="MIR repeat: matches 46. .204 of consensus"
repeat_region	9677..9718 /note="21 copies 2 mer tt 85% conserved"	repeat_region	32005..32260 /note="128 copies 2 mer ta 80% conserved"
repeat_region	9745..10036 /note="MER92C repeat: matches 272. .552 of consensus"	repeat_region	32903..33194 /note="AluJo repeat: matches 1. .293 of consensus"
repeat_region	10531..10751 /note="MER92B repeat: matches 2. .212 of consensus"	repeat_region	33677..34221 /note="MER68A repeat: matches 1. .568 of consensus"
repeat_region	10857..10969 /note="MIR repeat: matches 40. .154 of consensus"	repeat_region	34372..35093 /note="TIGGER1 repeat: matches 1650. .2308 of consensus"
repeat_region	11521..11613 /note="MIR repeat: matches 63. .156 of consensus"	repeat_region	35094..35333 /note="L1PB1 repeat: matches 5909. .6155 of consensus"
repeat_region	11629..12187 /note="L2 repeat: matches 1779. .2341 of consensus"	repeat_region	35334..35437 /note="TIGGER1 repeat: matches 2308. .2418 of consensus"
repeat_region	12438..12548 /note="L1MA9 repeat: matches 6163. .6270 of consensus"	repeat_region	35477..35878 /note="MLT1-INTERNAL repeat: matches 923. .1579 of consensus"
repeat_region	12975..13093 /note="MIR repeat: matches 98. .247 of consensus"	repeat_region	36082..36740 /note="MLT1-INTERNAL repeat: matches 201. .885 of consensus"
repeat_region	14083..14381 /note="AluSx repeat: matches 1. .293 of consensus"	repeat_region	36783..37121 /note="L2 repeat: matches 824. .1219 of consensus"
repeat_region	14906..15126 /note="L2 repeat: matches 2364. .2569 of consensus"	repeat_region	37158..37328 /note="L1M4 repeat: matches 4640. .4816 of consensus"
repeat_region	15142..15370 /note="MER4B repeat: matches 1. .215 of consensus"	repeat_region	39953..39978 /note="13 copies 2 mer ca 100% conserved"
repeat_region	15366..16028 /note="MER4D repeat: matches 359. .1017 of consensus"	repeat_region	40525..40566 /note="MIR repeat: matches 110. .151 of consensus"
repeat_region	16097..16211 /note="L2 repeat: matches 2631. .2749 of consensus"	repeat_region	44063..44377 /note="AluJo repeat: matches 1. .299 of consensus"
repeat_region	16199..16240 /note="L2 repeat: matches 2560. .2599 of consensus"	repeat_region	44451..45182 /note="MER21B repeat: matches 3. .790 of consensus"
repeat_region	18264..18349 /note="43 copies 2 mer ga 62% conserved"	repeat_region	45243..45477 /note="L1M4 repeat: matches 4078. .4321 of consensus"
repeat_region	19489..19797 /note="AluJb repeat: matches 1. .306 of consensus"	repeat_region	45515..45890 /note="L1PA7 repeat: matches 5769. .6143 of consensus"
repeat_region	19826..20044 /note="L1M3 repeat: matches 5220. .5419 of consensus"	repeat_region	45900..47783 /note="L1PA8 repeat: matches 4284. .6161 of consensus"
repeat_region	20045..20175 /note="FLAM_A repeat: matches 1. .131 of consensus"	repeat_region	47776..48115 /note="L1M4 repeat: matches 3680. .4029 of consensus"
repeat_region	20176..20249 /note="L1M3 repeat: matches 5419. .5496 of consensus"	repeat_region	48116..48477 /note="MER47A repeat: matches 1. .365 of consensus"
repeat_region	20270..20929 /note="L1MA5A repeat: matches 5447. .6104 of consensus"	repeat_region	48478..48554 /note="MER47A repeat: matches 1. .365 of consensus"
repeat_region	20930..21226 /note="AluSx repeat: matches 1. .297 of consensus"		
repeat_region	21227..21417 /note="L1MA5A repeat: matches 6104. .6294 of consensus"		
repeat_region	21418..21461 /note="L1M4 repeat: matches 4295. .4338 of consensus"		
repeat_region	21500..23928 /note="L1PA5 repeat: matches 3696. .6143 of consensus"		
repeat_region	24145..24341 /note="L1MEC repeat: matches 1820. .2012 of consensus"		

Query Match 21.58; Score 167.8; DB 9; Length 149749;
Best Local Similarity 70.8%; Pred. No. 4.4e-35;
Matches 288; Conservative 0; Mismatches 97; Indels 22; Gaps 4;
Qy 10 CTGTGAGCGCGACTCGGTCCTCGGTCGTGGAGTCTTTCTCGGTACTGACACGG 69
Db 76290 CTGTGAGCTGTGAGGCCAAGTCTCTAGAGTCTTTCTCTCTACTGAGATGACG 76349
Qy 70 CGGCTAGGTCCACAGGCAGATCCAACTGGGAGTTGAAGTGTGAGTGAAGAGGAA 129

closure of gap impossible.
The true right end of clone L21F12 is at 10842 in this sequence.
The true left end of clone L185B6 is at 1 in this sequence. L185B6
is from cosmid library LA04NC01 constructed at the Human Genome
Center, Los Alamos National Laboratory, NM 87545 under the auspices
of the U.S. Department of Energy. The library was constructed using
flow-sorted human chromosome 4 from a Hamster-Human hybrid cell
line (UV20HL21-27) containing human chromosomes 4, 8 and 21.
VECTOR: SCOS1
L185B6 is contained in a clone contig spanning
2Mb which is described in Baxendale et al, Nature Genetics 4 (1993
) 181- 186. See also Myers et al, Cytogenet Cell Genet. 66 (1994
) 218-230.

```
FEATURES
Source
1..20587
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/map="4p16.3"
/clone_lib="LA04NC01"
/cell_line="UV20HL21-27"
/clone="LA04NC01-185B6"
/complement(1370..1739)
/note="match: STS L00885"
2380..4066
/note="L1 element fragment"
4115..4400
/note="Alu repeat: matches 1..308 of consensus"
4427..4736
/note="L1 element fragment"
4778..5099
/note="L1 element fragment"
5203..5333
/note="MSTC element fragment"
5632..5721
/note="L1 element fragment"
5980..6714
/note="matches Trk-2h oncogene - 4laa of human RL7A_HUMAN
M36072"
/codon_start=1
/pseudo
6747..6778
/note="16 copies of 2 mer 100 % conserved"
10906..11082
/partial
/note="Alu repeat: matches 1..191 of consensus"
11083..11214
/partial
/note="Alu repeat: matches 175..308 of consensus"
11352..11506
/partial
/note="Alu repeat: matches 308..151 of consensus"
13249..13799
/note="putative CpG island"
15946..16238
/note="Alu repeat: matches 1..308 of consensus"
17069..17143
/note="MLTID element fragment"
17589..17689
/partial
/note="Alu repeat: matches 308..210 of consensus"
17695..17739
/partial
/note="Alu repeat: matches 170..114 of consensus"
18268..18369
/partial
/note="Alu repeat: matches 155..38 of consensus"
19668..19742
/note="MIR element fragment"
20313..20524
/note="4 copies of 53 mer 83 % conserved"
5439 a 5367 c 4601 g 5180 t
```

BASE COUNT
ORIGIN

```
Query Match      21.08; Score 164.2; DB 9; Length 20587;
Best Local Similarity 75.18; Pred. No. 3.9e-34;
Matches 247; Conservative 0; Mismatches 73; Indels 9; Gaps 3;

Oy 429 GAGCCCACTAATGGAGAGCCCAAAAAGCAAGACAGCAGCTGAAAGTCGGGATCCTACA 488
      ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
Db 18056 GAGCCCACTGATGAGCACCCTCAGCAAGAGGACACCACCTGAAAGTCGGGCTTT--A 17999

Oy 489 CCTGGGCAGCAGACAGAGAAGATCAGGTACAGTCCAGTCCAGTGCGGCAGCATGAA 548
      ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
Db 17998 CTTGCCAGGAGAGAGGAAGATCACGGTGTCTGTAGATTTCTGTCTTTCAGCAGGAA 17939

Oy 549 GGTGATCTGCAAGAGCTGCATCAGTCAACACCGGGGATAAATCTGGATTGTTGGTCCGG 608
      ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
Db 17938 GCTGATCTCCGGGAGCTGTCTC-----AAAGACTGGGATGAATGTGGAGATGG--TCCTG 17886

Oy 609 CGTCAAGGTGAAGATAATACCTTAAGAGGAACTGTAAAAATGCCAGAGCAGGTGAAGA 668
      || || || || || || || || || || || || || || || || || || || || || ||
Db 17885 ATGTCGGGGGAATATTCTGCCGCAATCAGAGCAATTTAAATTGCCAGAAGCAGGTGAAG 17826

Oy 669 GCACCCACAAGTTTAAATGAAGACAAAGCTGAACAAACCAAGCTGGTTTATATTAGATA 728
      ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
Db 17825 GCAACCCACAGGTTCCAGGAGAGCAAGCTGAACAAATGCAAACTGGTTTATATTAGATA 17766

Oy 729 TTTGACTTAAACTATCTCAATAAAGTTTT 757
      ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
Db 17765 CGTGACTTAAATAATCTCAATACAGTTTT 17737
```

```
RESULT 31
HS193G15/c
LOCUS
DEFINITION
        HS193G15                62493 bp      DNA                linear    PRI 15-SEP-1999
        Human DNA sequence from clone 193G15 on chromosome 4. Contains the
        5' part of the RGS12 gene for regulator of G-protein signalling 12,
        an RPL7A (60S Ribosomal Protein 7A (SURF3)) pseudogene, ESTs, an
        STS, GSSS and a putative CpG island, complete sequence.
ACCESSION
        AL117391
VERSION
        AL117391.1   GI:5911814
KEYWORDS
        HTG; 60S Ribosomal Protein 7A; CpG Island; G-protein; RGS12; RPL7A;
        SURF3.
SOURCE
        human.
ORGANISM
        Homo sapiens
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
        1 (bases 1 to 62493)
AUTHORS
        McMurray,A.
TITLE
        Direct Submission
JOURNAL
        Submitted (28-JUN-1999) Sanger Centre, Hinxton, Cambridgeshire,
        CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
        requests: clonerequest@sanger.ac.uk
COMMENT
        During sequence assembly data is compared from overlapping clones.
        Where differences are found these are annotated as variations
        together with a note of the overlapping clone name. Note that the
        variation annotation may not be found in the sequence submission
        corresponding to the overlapping clone, as we submit sequences with
        only a small overlap as described above.
        The following abbreviations are used to associate primary accession
        numbers given in the feature table with their source databases:
        EN: EMBL; SW: SWISSPROT; Tr: TREMBL
        IMPORTANT: This sequence is not the entire insert of clone 193G15.
        It may be shorter because we only sequence overlapping sections
        once, or longer because we arrange for a small overlap between
        neighbouring submissions.
        The true left end of clone 1B6 (L21F12) is at 6371 in this
        sequence. The true right end of clone 361H4 is at 104 in this
        sequence. This sequence has been finished according to sequencing
        criteria as follows. An attempt is made to resolve all sequencing
        problems, such as compressions and repeats, but not necessarily
        within known annotated human repeat sequence elements (e.g. Alu).
        Where the sequence is ambiguous, there is an annotation using the
        'unsure' feature key.
        This sequence was generated from part of bacterial clone contigs of
```


using the method described by Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from one male donor. For further details, see <http://bacpac.med.buffalo.edu/>. The clone is available from Genome Systems, Inc. (<http://www.genomesystems.com>).

VECTOR: pCYPAC2

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the right is RP5-848B16, 200 bp overlap. Actual start of this clone is at base position 1 of RP4-698G21; actual end is 109273 of RP4-698G21.

This clone contains STS SWSS3337 (NID:g1113739).

FEATURES

Location/Qualifiers

1..109469

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="7"

/map="7p21-p22"

/clone="RP4-698G21"

/clone_lib="RPC1-4"

386..420

/rpt_family="L2"

474..567

/rpt_family="L2"

586..962

/rpt_family="MaLR"

968..1041

/rpt_family="L2"

1045..1109

/rpt_family="MaLR"

1157..1410

/rpt_family="L2"

3108..3194

/rpt_family="MIR"

4869..11012

/rpt_family="L1"

11296..11387

/rpt_family="MIR"

11395..11702

/rpt_family="L1"

12129..12213

/rpt_family="MIR"

14481..14668

/rpt_family="MER1_type"

18027..18324

/rpt_family="L1"

21120..21421

/rpt_family="Alu"

21508..21627

/rpt_family="Alu"

21686..21804

/rpt_family="(TA)n"

22300..22367

/rpt_family="L1"

24376..24504

/rpt_family="L1"

24514..24578

/rpt_family="L2"

25630..26169

/rpt_family="MER2_type"

27828..28703

/rpt_family="L1"

29740..30022

/rpt_family="Alu"

30083..30286

/rpt_family="Alu"

30590..30891

/rpt_family="Alu"

34090..34390

/rpt_family="Alu"

34660..34976

/rpt_family="L1"

35070..35203

repeat_region
38192..39129
/rpt_family="L1"
repeat_region
39132..39429
/rpt_family="L1"
repeat_region
41307..41600
/rpt_family="Alu"
repeat_region
44883..45881
/rpt_family="L1"
repeat_region
46103..46316
/rpt_family="MIR"
repeat_region
49179..49480
/rpt_family="Alu"
repeat_region
50097..50184
/rpt_family="L2"
repeat_region
52842..52897
/rpt_family="L2"
repeat_region
53288..53667
/rpt_family="MaLR"
repeat_region
54194..54584
/rpt_family="MaLR"
repeat_region
55565..55867
/rpt_family="Alu"
repeat_region
57119..57245
/rpt_family="Retroviral"
repeat_region
57390..57660
/rpt_family="Retroviral"
repeat_region
57742..57865
/rpt_family="MIR"
misc_feature
complement(58417..58588)
/note="match to EST AA551305 (NID:g2321557) nk71b08.s1"
repeat_region
58589..58850
/rpt_family="Alu"
repeat_region
60889..61048
/rpt_family="MIR"
repeat_region
62270..62322
/rpt_family="(CA)n"
repeat_region
62528..63508
/rpt_family="L1"
repeat_region
63703..69828
/rpt_family="L1"
repeat_region
69839..75304
/rpt_family="L1"
repeat_region
76788..76902
/rpt_family="L2"
repeat_region
77385..77693
/rpt_family="Alu"
repeat_region
78108..78287
/rpt_family="MER1_type"
repeat_region
78771..79368
/rpt_family="L2"
repeat_region
79740..79832
/rpt_family="MIR"
repeat_region
82376..82765
/rpt_family="MaLR"
repeat_region
83415..83536
/rpt_family="Alu"
repeat_region
85229..85584
/rpt_family="MaLR"
repeat_region
86335..86595
/rpt_family="Alu"
repeat_region
86927..87104
/rpt_family="MIR"
repeat_region
88291..88397
/rpt_family="L1"
repeat_region
88394..88507
/rpt_family="L1"
repeat_region
88539..88650
/rpt_family="L2"
repeat_region
90033..90339

Query Match

Best Local Similarity 14.5%; Score 113.4; DB 9; Length 109469;
63.6%; Pred. No. 4.1e-20;

Matches 260; Conservative 0; Mismatches 131; Indels 18; Gaps 5;

Qy 1 CGCGCGGAGCTGTAGCCGCGA--CTCGGGTCCCTGAGCTGCGATCTTTTCGCGTA 58
|||||
Db 36964 CGCGCGGAGCTGTAGGGTGTGAGGTCGCGCTCTCTCTCGGACTTTTCTGTCTTA 37023

Qy 59 CTGAGACACGCGGAGTCCACAGCAGATCAACTGGGAGTTGAAGTGTGAGTGAGA 118
|||||
Db 37024 CTGAGACACGCGCGGTAGGTCCGAGGCCAGTCTTCCAGGGGTTGAAGTGTGAGT 37083

Qy 119 GTGAAGAGGAACACGAGGCTTCCGGAGGTTGTG-----TGTCTAGTGACTCAGAGTG 172
|||||
Db 37084 GTGAGGAGGACCCAGCGCGCTCTGGAGCATGGGCGAGCGTGTCCGTGCGCTCAGAGGC 37143

Qy 173 AGAAGGCCCTCGAAGTCGTCTCTCTCTATGCGGTGCCAGGCCCATGGACCTTCTTCTC 232
|||||
Db 37144 CGATGGGTCTCAAGGTATCGTCTCTCTCTCGTGGAGTGGCAGCCGATGGGCTTTGTT 37203

Qy 233 TCGTCACGGCCATAACTAGGAGGAAGGAGGCGGAGGAGTGGAG-GGGCTCAGGCGAAG 291
|||||
Db 37204 GTAGTGGGTGAGAACGAGGAAGGTGGCGCGCTAGTTGAGAAGGCTCAGTGAAG 37263

Qy 292 CTGGG----GTGCTGTGGGGGTATCCGAGTCCAGAGCACCTTGGAAACC-----CCGAC 342
|||||
Db 37264 ATGGGGTAAGTCTGCGGGGTCTGAGCAGGTATCGGAGTCCGTGGAAACCTCGACAGAG 37323

Qy 343 AGAGATTCTGCACTCCCGACAGCGGACCGAGGAGGCGGCATGAGC 391
|||||
Db 37324 GACGATTCTGCACTCTCCGTGCGAGCCAGGAAGGCGGGTGGCG 37372

RESULT 38
AF055473
LOCUS AF055473 Homo sapiens GAGE-8 mRNA, complete cds. PRI 01-MAY-2000
DEFINITION AF055473
ACCESSION AF055473
VERSION AF055473.1 GI:3511022
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 528)
De Backer,O., Arden,K.C., Boretti,M., Vantomme,V., De Smet,C.,
Czekay,S., Viars,C.S., De Plaen,E., Brasseur,F., Chomez,P., Van den
Eynde,B., Boon,T. and van der Bruggen,P.
Characterization of the GAGE genes that are expressed in various
human cancers and in normal testis
Cancer Res. 59 (13), 3157-3165 (1999)
JOURNAL 99323388
MEDLINE 10397259
PUBMED
REFERENCE 2 (bases 1 to 528)
De Backer,O.R.Y.
Direct Submission
TITLE Submitted (24-MAR-1998) Ludwig Institute for Cancer Research,
JOURNAL Brussels Branch, 74 av. Hippocrate, Brussels B-1200, Belgium
FEATURES
source
1..528
/organism="Homo sapiens"
/db_xref="taxon:9606"
73..423
CDS
/codon_start=1
/product="GAGE-8"
/protein_id="AAC33676.1"
/db_xref="GI:3511023"
/translation="MSWRGRSTYRPRRVRPEPMIGPMRPEQFSDSEVPATPERGE
PATQRQDPAAAQEGEDSGAGGPKPEADSQGHGQTCEDCEDGDGPEMDPPNPE
EVKTPERGEKQSQ"

BASE COUNT 158 a 112 c 151 g 107 t

Query Match 14.1%; Score 110; DB 9; Length 528;
Best Local Similarity 63.9%; Pred. No. 2.6e-19;
Matches 260; Conservative 0; Mismatches 131; Indels 18; Gaps 5;

Matches 216; Conservative 0; Mismatches 115; Indels 7; Gaps 3;

Qy 447 CCCCCAAAAGAAAGAACCCAGCAGCTGAAAGTCGGGATCCCTACACCTGGCGAGCAGAGAA 506
|||||
Db 190 CTTGAAGAAGGGGAAGCAGCAACTCAACGTCAAGGTCTGCGAGCTCTCAGGAG---GGA 246

Qy 507 GAAGATCAGGATACAGCTGAGATCCCAAGTCGCGAGCATGAAGGTGATCTGCAAGAGCTG 566
|||||
Db 247 GAGGATGAGGGAGCATCTGCAGGTCAAGGCCCAAGCCTGAAGCTGATAGCCAGGAACAG 306

Qy 567 CATCAGTCAAAACACCCGGGATAAATCTGGATTTGGGTTCGGCGTCAAGGTGAAGATAAT 626
|||||
Db 307 GGTCAACCCACAGACTGGGTGTGAGTGTGAAGAT-GGTCTGTATGGCGCAGAGATGGACCC 365

Qy 627 ACCTAAGAGGAGACACTGTAAATGCCAAGCAGGTGAAGAGCAACCAACCAAGTTTAAAT 686
|||||
Db 366 GCCAAATCAGAGGAGGTGAAGACGCTGAAGAGGTGAAGCAATCACAGTGTAAAA 425

Qy 687 GAAGACAAGTGAACACAGCAGCTGGTTTATATATAGATATTG---ACTTAAACTAT 743
|||||
Db 426 GAAGACAGTTGAATATGATGAGGCTGCTCTATGTTGAAATTTGTTCAATAAATCT 485

Qy 744 CTCAATAAGTTTTCGAGCTTTTCAACCAAAAAA 781
|||||
Db 486 CCAATAAGCTTTACAGGCTTCTGCAAGAAAAA 523

RESULT 39
AR028492
LOCUS AR028492 Sequence 18 from patent US 5858689.
DEFINITION AR028492
ACCESSION AR028492
VERSION AR028492.1 GI:5940465
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 539)
van der Bruggen,P., van den Eynde,B., DeBacker,O. and
Boon-Failleur,T.
Isolated peptides derived from the gage tumor rejection antigen
precursor and uses thereof
Patent: US 5858689-A 18 12-JAN-1999;
JOURNAL Location/Qualifiers
FEATURES
source
1..539
/organism="unknown"
BASE COUNT 158 a 113 c 157 g 111 t

Query Match 14.1%; Score 110; DB 6; Length 539;
Best Local Similarity 63.9%; Pred. No. 2.6e-19;
Matches 216; Conservative 0; Mismatches 115; Indels 7; Gaps 3;

Qy 447 CCCCCAAAAGAAAGAACCCAGCAGCTGAAAGTCGGGATCCTACACCTGGCGAGCAGAGAA 506
|||||
Db 202 CCTGAAGAAGGGGAAGCAGCAACTCAACGTCAAGGTCTGCGAGCTCTCAGGAG---GGA 258

Qy 507 GAAGATCAGGATACAGCTGAGATCCCAAGTCGCGACATGAAGGTGATCTGCAAGAGCTG 566
|||||
Db 259 GAGGATGAGGAGGATCTGCAGGTCAAGGGCCGAAGCCTGAAGCTGATAGCCAGGAACAG 318

Qy 567 CATCAGTCAAAACACCCGGGATAAATCTGGATTTGGGTTCGGCGTCAAGGTGAAGATAAT 626
|||||
Db 319 GGTCAACCCACAGACTGGGTGTGAGTGTGAAGAT-GGTCTGTATGGCGAGGAGGTGGACCC 377

Qy 627 ACCTAAGAGAGAACACTGTAAATGCCAAGAGCAGGTGAAGAGCAACCAACCAAGTTTAAAT 686
|||||
Db 378 GCCAAATCCAGAGGAGGTGAAGACGCTGAAGAGGTGAAGCAATCACAGTGTAAAA 437

Qy 687 GAAGACAAGCTGAACCAACCAAGCTGTTTATATATAGATATTG---ACTTAAACTAT 743
|||||
Db 438 GAAGACAGTTGAATATGAGGCTGCTCTATGTTGGAATTTGTTCAATAAATCT 497

Thu Oct 17 09:19:16 2002

83	18	2.3	1396	22	AAH17151	Human cDNA sequenc	c 156	17	2.2	474	22	AAK62159	Human immune/haema
84	18	2.3	1471	22	AAI59131	Human polynucleoti	c 157	17	2.2	475	22	ABA25109	Probe #3575 for qe
85	18	2.3	1572	22	AAF27283	Human atonal homol	c 158	17	2.2	475	22	AAK03627	Human brain expres
86	18	2.3	1832	21	AAK36377	Arabidopsis thalia	c 159	17	2.2	475	22	AAK29086	Human bone marrow
87	18	2.3	1836	21	AAK50584	Arabidopsis thalia	c 160	17	2.2	475	22	AAI03552	Probe #3543 used t
88	18	2.3	1853	21	AAK78405	Human secreted pro	c 161	17	2.2	536	22	AAH10605	Human cDNA clone (
89	18	2.3	1897	18	AAZ71258	Human secreted pro	c 162	17	2.2	590	22	AAH34745	Human colon cancer
90	18	2.3	1897	18	AAZ71258	Proviral oncogene	c 163	17	2.2	612	20	AAZ24864	Human secreted pro
91	18	2.3	1967	22	AAK00836	Nucleotide sequenc	c 164	17	2.2	620	21	AAK54242	Arabidopsis thalia
92	18	2.3	2394	22	AAK85095	Human cDNA clone H	c 165	17	2.2	684	23	AAK43832	Arabidopsis thalia
93	18	2.3	2714	23	ABL22182	Atherosclerosis-as	c 166	17	2.2	708	20	AAI18952	Partial human PIGR
94	18	2.3	3261	20	AAV64855	Drosophila melanog	c 167	17	2.2	708	21	AAK54365	Arabidopsis thalia
95	18	2.3	3410	22	AAK53375	Human LFA-1 alpha	c 168	17	2.2	709	21	AAK53473	Arabidopsis thalia
96	18	2.3	3510	22	AAK64853	Human polynucleoti	c 170	17	2.2	731	21	AAK43954	Arabidopsis thalia
97	18	2.3	3585	17	AAZ71258	Human LFA-1 alpha	c 171	17	2.2	744	24	AAI71487	Human pax protein
98	18	2.3	3650	23	AAK72651	Human LFA-1 alpha	c 172	17	2.2	752	20	AAK98945	Human validated ca
99	18	2.3	3650	23	AAK72651	Human LFA-1 alpha	c 173	17	2.2	752	20	AAK98945	Human validated ca
100	18	2.3	4980	22	AAK66684	Mouse Rim2-encodin	c 174	17	2.2	763	22	AAI96392	Human neuroblastom
101	18	2.3	5049	22	AAK66684	Human low adenosin	c 175	17	2.2	777	20	AAK99046	Human validated ca
102	18	2.3	5133	21	AAK54333	Human LFA-1 alpha	c 176	17	2.2	789	20	AAK08693	Novel nucleotide s
103	18	2.3	5133	21	AAK54333	Human LFA-1 alpha	c 177	17	2.2	789	20	AAK08693	CPE 114 coding seq
104	18	2.3	5133	21	AAK54333	Human LFA-1 alpha	c 178	17	2.2	789	20	AAK08693	Arabidopsis thalia
105	18	2.3	5133	21	AAK54333	Human LFA-1 alpha	c 179	17	2.2	801	23	AAK55745	Streptococcus pneu
106	18	2.3	5680	22	AAK46768	Tumour suppressor	c 180	17	2.2	801	23	AAK55745	Streptococcus pneu
107	18	2.3	6012	24	AAK61274	Human immune syste	c 181	17	2.2	813	22	AAH90915	2CFE 114 coding se
108	18	2.3	6287	22	ABA20058	Human gene regulat	c 182	17	2.2	814	21	AAK43817	Arabidopsis thalia
109	18	2.3	7223	22	AAK85189	Human immune syste	c 183	17	2.2	858	22	AAK71698	B. subtilis panc g
110	18	2.3	7223	22	AAK85189	Human immune syste	c 184	17	2.2	864	23	AAK71698	DNA encoding novel
111	18	2.3	8064	24	ABL33386	Human nervous syst	c 185	17	2.2	1007	20	AAV08829	Gene No. 19 encodi
112	18	2.3	8064	24	ABL33386	Human immune/haema	c 186	17	2.2	1020	20	AAK81240	Genomic sequence #
113	18	2.3	9646	24	ABL33386	Human immune/haema	c 187	17	2.2	1020	20	AAK81240	Human digestive sy
114	18	2.3	10820	24	ABL33386	Human immune syste	c 188	17	2.2	1079	22	AAK71698	Human leukocyte ce
115	18	2.3	11015	24	AAZ22335	Human immune syste	c 189	17	2.2	1079	22	AAK71698	Human diagnostic a
116	18	2.3	11726	24	ABL34062	Mouse ischaemic co	c 190	17	2.2	1113	19	AAV49553	Human PRO1565 (UNQ
117	18	2.3	12054	24	ABL33178	Chemically treated	c 191	17	2.2	1113	19	AAV49553	Human PRO1565 (UNQ
118	18	2.3	12537	17	AAZ11705	Human immune syste	c 192	17	2.2	1113	19	AAV49553	Human PRO1565 (UNQ
119	18	2.3	13102	22	ABA19342	Human immune syste	c 193	17	2.2	1113	19	AAV49553	Primer #94 used in
120	18	2.3	13102	22	ABA19342	Lymphocyte specifi	c 194	17	2.2	1197	21	AAK37112	Human angiotensin
121	18	2.3	13102	22	ABA19342	Human nervous syst	c 195	17	2.2	1197	21	AAK37112	Human chondromodul
122	18	2.3	13102	22	ABA19342	Drosophila melanog	c 196	17	2.2	1197	21	AAK37112	DNA encoding a hum
123	18	2.3	13102	22	ABA19342	Chemically pretrea	c 197	17	2.2	1206	22	AAH26446	Human secreted pro
124	18	2.3	13102	22	ABA19342	Human immune/haema	c 198	17	2.2	1206	22	AAH26446	Human colon cancer
125	18	2.3	13102	22	ABA19342	Human immune/haema	c 199	17	2.2	1228	20	AAK51746	Arabidopsis thalia
126	18	2.3	13102	22	ABA19342	Human immune/haema	c 200	17	2.2	1228	20	AAK51746	Arabidopsis thalia
127	18	2.3	13102	22	ABA19342	Human phosphatase	c 201	17	2.2	1228	20	AAK51746	Breast cancer prot
128	18	2.3	13102	22	ABA19342	Human genome from	c 202	17	2.2	1228	20	AAK51746	Human 50 kDa dyltr
129	18	2.3	13102	22	ABA19342	Human secreted pro	c 203	17	2.2	1228	20	AAK51746	Human (50 kDa) dyl
130	18	2.3	13102	22	ABA19342	Human secreted pro	c 204	17	2.2	1228	20	AAK51746	Human secreted pro
131	17	2.2	13102	22	ABA19342	Human secreted pro	c 205	17	2.2	1228	20	AAK51746	Human polynucleoti
132	17	2.2	13102	22	ABA19342	Human secreted pro	c 206	17	2.2	1228	20	AAK51746	Human secreted pro
133	17	2.2	13102	22	ABA19342	Human secreted pro	c 207	17	2.2	1228	20	AAK51746	Human polynucleoti
134	17	2.2	13102	22	ABA19342	Human secreted pro	c 208	17	2.2	1228	20	AAK51746	Human polynucleoti
135	17	2.2	13102	22	ABA19342	Human secreted pro	c 209	17	2.2	1228	20	AAK51746	Human polynucleoti
136	17	2.2	13102	22	ABA19342	Human secreted pro	c 210	17	2.2	1228	20	AAK51746	Human polynucleoti
137	17	2.2	13102	22	ABA19342	Human secreted pro	c 211	17	2.2	1228	20	AAK51746	Human polynucleoti
138	17	2.2	13102	22	ABA19342	Human secreted pro	c 212	17	2.2	1228	20	AAK51746	Human polynucleoti
139	17	2.2	13102	22	ABA19342	Human secreted pro	c 213	17	2.2	1228	20	AAK51746	Human polynucleoti
140	17	2.2	13102	22	ABA19342	Human secreted pro	c 214	17	2.2	1228	20	AAK51746	Human polynucleoti
141	17	2.2	13102	22	ABA19342	Human secreted pro	c 215	17	2.2	1228	20	AAK51746	Human polynucleoti
142	17	2.2	13102	22	ABA19342	Human secreted pro	c 216	17	2.2	1228	20	AAK51746	Human polynucleoti
143	17	2.2	13102	22	ABA19342	Human secreted pro	c 217	17	2.2	1228	20	AAK51746	Human polynucleoti
144	17	2.2	13102	22	ABA19342	Human secreted pro	c 218	17	2.2	1228	20	AAK51746	Human polynucleoti
145	17	2.2	13102	22	ABA19342	Human secreted pro	c 219	17	2.2	1228	20	AAK51746	Human polynucleoti
146	17	2.2	13102	22	ABA19342	Human secreted pro	c 220	17	2.2	1228	20	AAK51746	Human polynucleoti
147	17	2.2	13102	22	ABA19342	Human secreted pro	c 221	17	2.2	1228	20	AAK51746	Human polynucleoti
148	17	2.2	13102	22	ABA19342	Human secreted pro	c 222	17	2.2	1228	20	AAK51746	Human polynucleoti
149	17	2.2	13102	22	ABA19342	Human secreted pro	c 223	17	2.2	1228	20	AAK51746	Human polynucleoti
150	17	2.2	13102	22	ABA19342	Human secreted pro	c 224	17	2.2	1228	20	AAK51746	Human polynucleoti
151	17	2.2	13102	22	ABA19342	Human secreted pro	c 225	17	2.2	1228	20	AAK51746	Human polynucleoti
152	17	2.2	13102	22	ABA19342	Human secreted pro	c 226	17	2.2	1228	20	AAK51746	Human polynucleoti
153	17	2.2	13102	22	ABA19342	Human secreted pro	c 227	17	2.2	1228	20	AAK51746	Human polynucleoti
154	17	2.2	13102	22	ABA19342	Human secreted pro	c 228	17	2.2	1228	20	AAK51746	Human polynucleoti
155	17	2.2	13102	22	ABA19342	Human secreted pro	c 229	17	2.2	1228	20	AAK51746	Human polynucleoti

17	375	2.2	28818	22	AAL35901	Human musculoskele	c 448	16	2.0	255	22	AAH82245	Rat differential t
17	376	2.2	31056	22	AAL03450	Human reproductive	c 449	16	2.0	258	22	AAK57059	Human immune/haema
17	377	2.2	31969	23	ABL07768	Drosophila melanog	c 450	16	2.0	264	21	AAH87519	Rat hepatocyte car
17	378	2.2	32236	22	AK91250	Human digestive sy	c 451	16	2.0	267	21	AAH02146	Human secreted pro
17	379	2.2	32236	22	AK91250	Human immune/haema	c 452	16	2.0	270	14	AAO60682	Human brain expres
17	380	2.2	37925	22	AAK86321	Human immune/haema	c 453	16	2.0	271	21	AAH11751	Aspergillus niger
17	381	2.2	37925	22	AAK86321	Human immune/haema	c 454	16	2.0	281	16	AAO99324	Mycobacterium kans
17	382	2.2	44602	22	AAK72372	Human immune/haema	c 455	16	2.0	283	20	AAV89992	EST clone CW762.
17	383	2.2	46954	23	ABL16830	Drosophila melanog	c 456	16	2.0	288	21	AAA42875	Human secreted exp
17	384	2.2	51402	22	AAK72363	Human immune/haema	c 457	16	2.0	288	21	AAA42875	Human breast cance
17	385	2.2	54863	22	AAK86025	Human immune/haema	c 458	16	2.0	289	22	AAH24984	Human breast cance
17	386	2.2	56743	22	AAK86026	Human immune/haema	c 459	16	2.0	289	22	AAH88396	Human polynucleoti
17	387	2.2	57728	22	AAK68202	Human immune/haema	c 460	16	2.0	295	22	AAH59438	Human KCC2 genomic
17	388	2.2	66685	22	AAK81760	Human immune/haema	c 461	16	2.0	297	22	AAH88715	Human digestive sy
17	389	2.2	66685	22	AAK81760	Human immune/haema	c 462	16	2.0	297	22	AAH88715	Human liver associ
17	390	2.2	73334	24	ABL34124	Human 9p11 chromos	c 463	16	2.0	300	20	AAH31749	Human gene express
17	391	2.2	73947	23	ABL07230	Human nibrin DNA	c 464	16	2.0	300	20	AAH31749	Human gene express
17	392	2.2	87350	18	AAH83003	Human genomic syste	c 465	16	2.0	300	20	AAH31749	Human gene express
17	393	2.2	162450	21	AAH83003	Drosophila melanog	c 466	16	2.0	305	22	AAH72363	Human cervical can
17	394	2.2	163319	21	AAH83003	Human WRN genomic	c 467	16	2.0	305	22	AAH72363	Human cervical can
17	395	2.2	611590	21	AAH22303	Retinoblastoma bin	c 468	16	2.0	311	18	AAH91705	Gene 056 fragment
17	396	2.2	1230025	20	AAH91990	Arabidopsis thalia	c 469	16	2.0	311	18	AAH91705	Gene 056 fragment
17	397	2.2	1230025	20	AAH91990	Arabidopsis thalia	c 470	16	2.0	311	18	AAH91705	Gene 056 fragment
17	398	2.2	1230025	20	AAH91990	Arabidopsis thalia	c 471	16	2.0	311	18	AAH91705	Gene 056 fragment
17	399	2.2	1230025	20	AAH91990	Arabidopsis thalia	c 472	16	2.0	311	18	AAH91705	Gene 056 fragment
17	400	2.2	1230025	20	AAH91990	Arabidopsis thalia	c 473	16	2.0	311	18	AAH91705	Gene 056 fragment
17	401	2.2	1230025	20	AAH91990	Arabidopsis thalia	c 474	16	2.0	311	18	AAH91705	Gene 056 fragment
17	402	2.2	1230025	20	AAH91990	Arabidopsis thalia	c 475	16	2.0	311	18	AAH91705	Gene 056 fragment
17	403	2.2	1230025	20	AAH91990	Arabidopsis thalia	c 476	16	2.0	311	18	AAH91705	Gene 056 fragment
17	404	2.2	1230025	20	AAH91990	Arabidopsis thalia	c 477	16	2.0	311	18	AAH91705	Gene 056 fragment
17	405	2.2	1230025	20	AAH91990	Arabidopsis thalia	c 478	16	2.0	311	18	AAH91705	Gene 056 fragment
17	406	2.2	1230025	20	AAH91990	Arabidopsis thalia	c 479	16	2.0	311	18	AAH91705	Gene 05

c 521	16	2.0	424	22	AAH98722	Murine EST-derived	594	16	2.0	579	22	ABA64267	Human foetal liver
c 522	16	2.0	426	22	AAI84618	Human polynucleoti	595	16	2.0	579	22	AAK1407	Probe #9873 for ge
c 523	16	2.0	427	22	AAI90659	Human polynucleoti	596	16	2.0	579	22	AAK12737	Human brain expres
c 524	16	2.0	428	20	AAV87921	EST clone FF97. H	597	16	2.0	579	22	AAK38456	Human bone marrow
c 525	16	2.0	431	22	AAI81704	Human polynucleoti	598	16	2.0	579	22	AAI19256	Probe #1312 used
c 526	16	2.0	432	22	AAK72723	Human immune/haema	599	16	2.0	579	22	AAI44426	Arabidopsis thalia
c 527	16	2.0	437	22	AAI88950	Human polynucleoti	600	16	2.0	584	21	AAC40423	Human excretory re
c 528	16	2.0	438	22	AAI10590	Human breast cance	c 601	16	2.0	584	22	AAI98605	Human kidney relat
c 529	16	2.0	439	22	AAK78629	Human immune/haema	c 602	16	2.0	584	22	AAI63001	CDNA sequence #614
c 530	16	2.0	440	22	AAI80553	Human polynucleoti	c 603	16	2.0	595	24	AAAG6817	Human foetal liver
c 531	16	2.0	454	21	AAH82087	N. meningitidis pa	c 604	16	2.0	596	22	AAK09102	Human brain expres
c 532	16	2.0	455	22	AAI92050	Human polynucleoti	c 605	16	2.0	596	22	AAK34992	Human bone marrow
c 533	16	2.0	460	22	AAH12874	Human cDNA clone (c 606	16	2.0	596	22	AAI40709	Human lung tumour
c 534	16	2.0	469	22	AAI82196	Human polynucleoti	c 607	16	2.0	599	24	ABK16292	Yeast YLR110C prom
c 535	16	2.0	470	21	AAC40259	Arabidopsis thalia	c 608	16	2.0	600	23	AAAD7502	DNA encoding novel
c 536	16	2.0	472	22	ABA59453	Human foetal liver	c 609	16	2.0	604	22	AAAD7502	Human breast cance
c 537	16	2.0	472	22	ABA28109	Probe #6575 for ge	c 610	16	2.0	604	22	AAI19264	Human breast cance
c 538	16	2.0	472	22	AAK07691	Human brain expres	c 611	16	2.0	604	22	AAI25033	Nucleotide sequenc
c 539	16	2.0	472	22	AAK33505	Human bone marrow	c 612	16	2.0	615	22	AAH42839	Arabidopsis thalia
c 540	16	2.0	472	22	AAI16475	Probe #6408 for ge	c 613	16	2.0	615	22	AAH42839	Human secreted pro
c 541	16	2.0	472	22	AAI39275	Probe #7961 used t	c 614	16	2.0	619	21	AAC59908	Human cDNA clone (
c 542	16	2.0	473	22	AAH44844	Human contig polyn	c 615	16	2.0	620	21	AAH12855	DNA encoding Hmhk-
c 543	16	2.0	476	22	AAH83118	Human ovarian tumo	c 616	16	2.0	624	22	AAH12855	Human secreted pro
c 544	16	2.0	477	21	AAC06385	Human secreted pro	c 617	16	2.0	627	17	AAH40734	Human cDNA clone (
c 545	16	2.0	478	22	AAK79254	Human immune/haema	c 618	16	2.0	636	21	AAC59368	DNA encoding Hmhk-
c 546	16	2.0	479	22	AAH7669	Peppermint plant o	c 619	16	2.0	640	22	AAH29046	Human secreted for
c 547	16	2.0	484	10	AAH92928	CDNA sequence from	c 620	16	2.0	640	22	AAH29121	CDNA encoding for
c 548	16	2.0	484	10	AAH92928	Barley Clone F. H	c 621	16	2.0	640	22	AAH29121	Human immune/haema
c 549	16	2.0	493	16	AAQ97352	Mycobacterium kans	c 622	16	2.0	642	22	AAK64676	Mycobacterium tube
c 550	16	2.0	493	16	AAQ97352	Yeast promoter YLR	c 623	16	2.0	645	24	AAH34407	Human G protein-co
c 551	16	2.0	494	22	AAH07474	Human polynucleoti	c 624	16	2.0	651	22	AAH34407	Human colon cancer
c 552	16	2.0	496	22	AAH87743	Secreted protein E	c 625	16	2.0	653	19	AAV52938	Pig FPCD antigen e
c 553	16	2.0	506	20	AAH24081	Human ovarian can	c 626	16	2.0	653	22	AAK86438	Human immune/haema
c 554	16	2.0	507	22	AAH20623	Human immune/haema	c 627	16	2.0	654	22	AAK64037	Human immune/haema
c 555	16	2.0	508	22	AAK62124	Human signal pepti	c 628	16	2.0	662	22	AAK88730	Human digestive sy
c 556	16	2.0	518	21	AAH98140	Human immune/haema	c 629	16	2.0	662	22	AAK88730	Human liver associ
c 557	16	2.0	519	22	AAK81405	Human bone marrow	c 630	16	2.0	670	22	AAK86439	Human immune/haema
c 558	16	2.0	520	22	AAK36414	Human bone marrow	c 631	16	2.0	677	23	AAH68439	DNA encoding novel
c 559	16	2.0	524	16	AAH05515	Human neurogenic d	c 632	16	2.0	683	23	AAH84127	DNA encoding novel
c 560	16	2.0	524	18	AAH74889	Human neurogenic d	c 633	16	2.0	685	22	AAH99850	Human protein enco
c 561	16	2.0	524	19	AAV42930	Partial DNA sequen	c 634	16	2.0	689	21	AAC43438	Arabidopsis thalia
c 562	16	2.0	530	21	AAH29475	Drosophila melanog	c 635	16	2.0	694	8	AAH71218	Sequence of the 5'
c 563	16	2.0	532	21	AAH40641	Arabidopsis thalia	c 636	16	2.0	694	8	AAH70210	Sequence of 5' fla
c 564	16	2.0	533	21	AAH27129	Human inflammation	c 637	16	2.0	697	22	ABA14515	Human nervous syst
c 565	16	2.0	533	9	AAH80239	Region between loc	c 638	16	2.0	697	22	ABA14515	Human nervous syst
c 566	16	2.0	533	21	AAH78489	Plant SDF polynuc	c 639	16	2.0	700	22	AAH92015	Human inflammatory
c 567	16	2.0	539	22	AAH89209	Human polynucleoti	c 640	16	2.0	700	22	AAH92015	Human inflammatory
c 568	16	2.0	539	22	AAK57950	Human immune/haema	c 641	16	2.0	700	22	AAH92819	Human inflammatory
c 569	16	2.0	542	22	ABA62828	Human foetal liver	c 642	16	2.0	700	22	AAH92819	Human inflammatory
c 570	16	2.0	542	22	ABA30112	Probe #8578 for ge	c 643	16	2.0	705	22	AAH04500	Human cDNA clone (
c 571	16	2.0	542	22	AAK11228	Human brain expres	c 644	16	2.0	705	22	ABL24197	Drosophila melanog
c 572	16	2.0	542	22	AAK37026	Human bone marrow	c 645	16	2.0	705	23	AAH24197	Human neuroblastom
c 573	16	2.0	542	22	AAH17873	Probe #7806 for ge	c 646	16	2.0	707	22	AAI95586	Arabidopsis thalia
c 574	16	2.0	542	22	AAI12846	Probe #11532 used	c 647	16	2.0	716	21	AAC39966	Human cervical can
c 575	16	2.0	543	22	AAH25687	H. pylori HPC172 e	c 648	16	2.0	722	22	AAH70534	Human breast cance
c 576	16	2.0	544	22	AAH50674	Human tumour assoc	c 649	16	2.0	732	22	AAH22026	Human immune/haema
c 577	16	2.0	546	22	ABA20255	Human nervous syst	c 650	16	2.0	741	22	AAK82469	Human breast cance
c 578	16	2.0	546	22	AAH32970	DNA encoding CARDI	c 651	16	2.0	744	21	AAC49543	Arabidopsis thalia
c 579	16	2.0	550	22	AAH94910	Cat flea hindgut a	c 652	16	2.0	748	21	AAC49543	Arabidopsis thalia
c 580	16	2.0	553	19	AAV34276	Human secreted pro	c 653	16	2.0	751	19	AAV53371	DNA encoding a Sta
c 581	16	2.0	553	23	AAH64348	DNA encoding novel	c 654	16	2.0	752	22	AAI21621	Human breast cance
c 582	16	2.0	555	23	ABL21791	Drosophila melanog	c 655	16	2.0	754	22	AAI23289	Human gene express
c 583	16	2.0	557	19	AAV34223	Human secreted pro	c 656	16	2.0	755	20	AAH217396	Human neuroblastom
c 584	16	2.0	561	22	AAC89051	Mannanase mannan b	c 657	16	2.0	755	22	AAI96715	Human breast cance
c 585	16	2.0	571	22	ABA62786	Human foetal liver	c 658	16	2.0	758	22	AAI23300	Human neuroblastom
c 586	16	2.0	571	22	ABA30079	Probe #8545 for ge	c 659	16	2.0	761	22	AAI96070	DNA encoding novel
c 587	16	2.0	571	22	AAI35685	Human musculoskele	c 660	16	2.0	772	23	AAH78469	Human neuroblastom
c 588	16	2.0	571	22	AAK11184	Human brain expres	c 661	16	2.0	775	22	AAI95900	Human immune/haema
c 589	16	2.0	571	22	AAK36991	Human bone marrow	c 662	16	2.0	775	22	AAH29120	Human secreted pro
c 590	16	2.0	571	22	AAH17833	Probe #7766 for ge	c 663	16	2.0	776	20	AAH29120	Human prostate tum
c 591	16	2.0	571	22	AAI42803	Probe #11489 used	c 664	16	2.0	777	20	AAH29120	Human secreted pro
c 592	16	2.0	572	22	AAI41139	Probe #9825 used t	c 665	16	2.0	803	23	AAH04317	Human secreted pro
c 593	16	2.0	574	22	AAI19901	Human breast cance	c 666	16	2.0				

c 813	16	2.0	1790	21	AAC341196	Arabidopsis thalia	886	16	2.0	2533	22	AAC88111	Human FLEXHT-42 nu
814	16	2.0	1794	23	AAS70163	DNA encoding novel	887	16	2.0	2544	22	AAC17822	Human cDNA sequenc
815	16	2.0	1824	21	AAT771128	Human ORFX ORF2683	c 888	16	2.0	2558	22	AAF55249	Nucleotide sequenc
816	16	2.0	1842	23	AAS72632	DNA encoding novel	c 889	16	2.0	2588	22	AAH73967	Human reducing aye
817	16	2.0	1844	19	AAV24018	Human BHF1 coding	890	16	2.0	2625	21	AAH16658	Human secreted pro
818	16	2.0	1845	15	AQ298400	Canine betal-adren	891	16	2.0	2627	21	AAH70077	Human ovarian carc
819	16	2.0	1875	15	AAQ58758	Rat glucagon recep	892	16	2.0	2627	21	AAH55681	Human ovarian tumo
820	16	2.0	1907	22	AAH19183	Human secreted pro	c 893	16	2.0	2630	23	AA569483	DNA encoding novel
c 821	16	2.0	1922	15	AAQ66639	Mouse Activin rece	894	16	2.0	2633	23	AA569279	Human secreted pro
c 822	16	2.0	1922	20	AAQ209846	Human MALK-1 clone	895	16	2.0	2635	22	ABA84521	Human secreted pro
823	16	2.0	1926	22	AAH19220	Human secreted pro	896	16	2.0	2658	23	ABL21790	Drosophila melanog
824	16	2.0	1944	21	AAAG69015	Bacteriophage 44AH	c 897	16	2.0	2665	22	AA501479	Human secreted pro
825	16	2.0	1954	22	AAQ66894	Human EXMAD-5 codi	898	16	2.0	2674	21	AA934443	GFP-HSC70 fusion g
c 826	16	2.0	1979	22	AAH18361	Human cDNA sequenc	899	16	2.0	2682	21	AA488863	Human dentin matril
c 827	16	2.0	1980	20	AA568808	Human leukaemia-as	900	16	2.0	2704	23	ABL13675	Drosophila melanog
c 828	16	2.0	1986	21	AAQ95489	Human secreted pro	c 901	16	2.0	2705	23	ABL124196	Drosophila melanog
c 829	16	2.0	2027	23	AA590594	DNA encoding novel	902	16	2.0	2719	22	AAH13886	Human cDNA sequenc
c 830	16	2.0	2054	21	AAQ59693	Human secreted pro	c 903	16	2.0	2720	23	ABL21236	Drosophila melanog
c 831	16	2.0	2058	14	AAQ41227	Clone GP46 encodin	904	16	2.0	2730	22	AA84293	Human EXCS encodin
c 832	16	2.0	2068	22	AAK81547	Human immune/haema	905	16	2.0	2732	20	AA84293	Mouse pheromone re
c 833	16	2.0	2081	21	AAK76582	Human ORFX ORF2137	906	16	2.0	2732	20	AA84293	Human cDNA sequenc
834	16	2.0	2089	22	AA557139	cDNA encoding Dros	907	16	2.0	2756	22	AAH17562	Human cDNA sequenc
835	16	2.0	2089	23	ABL13921	Drosophila melanog	c 908	16	2.0	2771	22	AAH17350	Human PRO4993 nucl
836	16	2.0	2089	24	ABI99790	Mouse ischaemic co	909	16	2.0	2840	21	AA878596	Drosophila melanog
c 837	16	2.0	2094	21	AAQ99018	Human pancreatic c	910	16	2.0	2865	23	ABL22130	Nucleotide sequenc
c 838	16	2.0	2101	17	AAAT11743	Human kynurenine ami	c 911	16	2.0	2895	22	AAF55248	Nucleotide sequenc
c 839	16	2.0	2114	20	AAK19958	Rat Tango-76 encod	912	16	2.0	2918	22	AAF55252	Human secreted pro
c 840	16	2.0	2118	21	AAK51464	Arabidopsis thalia	c 913	16	2.0	2923	20	AAK04326	Human stomach canc
841	16	2.0	2127	22	AAK85819	HOAT1 DNA. Homo s	c 914	16	2.0	2925	22	AAH16302	Human cDNA sequenc
842	16	2.0	2174	21	AA53823	Sequence encoding	c 915	16	2.0	2925	22	AAK86215	Human immune/haema
c 843	16	2.0	2198	20	AAK22779	Rat SOCS17 cDNA.	c 916	16	2.0	2957	22	AAK86215	Porcine BAC-PIGF2-
c 844	16	2.0	2198	20	AAK22773	Rat SOCS17 cDNA.	c 917	16	2.0	3000	21	AA65523	Nucleotide sequenc
845	16	2.0	2198	24	ABA03615	Human 34P3D7 cDNA.	c 918	16	2.0	3014	22	AAF55247	Drosophila melanog
846	16	2.0	2198	21	ABA03616	Human 34P3D7 mRNA.	c 919	16	2.0	3029	23	ABL23754	Nucleotide sequenc
847	16	2.0	2201	21	AAZ43777	Human fetal brain	c 920	16	2.0	3043	22	AAF55253	Drosophila melanog
848	16	2.0	2210	23	ABL24954	Drosophila melanog	c 921	16	2.0	3045	23	ABL19459	Drosophila melanog
849	16	2.0	2233	22	AA26418	Pseudomonas sp hea	922	16	2.0	3089	23	ABL04273	Nucleotide sequenc
c 850	16	2.0	2239	23	ABL11831	Drosophila melanog	c 923	16	2.0	3092	22	AAF55257	Arabidopsis thalia
c 851	16	2.0	2243	22	AAH73362	Human cervical can	c 924	16	2.0	3129	21	AAK43286	Nucleotide sequenc
c 852	16	2.0	2253	21	AAQ99127	Human pancreatic c	c 925	16	2.0	3136	22	AAF55250	Drosophila melanog
c 853	16	2.0	2275	22	AAK31140	Human diagnostic a	926	16	2.0	3136	23	ABL19748	Drosophila melanog
c 854	16	2.0	2285	17	AAAT13426	Humicola insolens	c 927	16	2.0	3152	23	ABL13448	Drosophila melanog
c 855	16	2.0	2285	19	AAV19376	Humicola insolens	c 928	16	2.0	3159	22	AAF55251	Nucleotide sequenc
856	16	2.0	2290	22	AA529606	Human endocrine po	c 929	16	2.0	3170	22	AAF55251	Drosophila melanog
857	16	2.0	2297	23	AA582982	DNA encoding novel	c 930	16	2.0	3223	23	ABL08588	Human c-myc oncoge
858	16	2.0	2303	18	AAAT47338	Murine development	c 931	16	2.0	3225	19	AAV20467	Human immune syste
859	16	2.0	2303	20	AAK18506	Murine del-1 encod	c 932	16	2.0	3231	24	ABL34288	Nucleotide sequenc
860	16	2.0	2308	18	AAAT47339	Murine del-1 trunc	c 933	16	2.0	3253	22	AAF55254	Nucleotide sequenc
c 861	16	2.0	2308	20	AAK18507	Truncated murine d	c 934	16	2.0	3253	22	AAF55256	Nucleotide sequenc
c 862	16	2.0	2309	23	ABL24854	Drosophila melanog	c 935	16	2.0	3254	22	AAF55255	Nucleotide sequenc
c 863	16	2.0	2314	22	AAK76171	Human immune/haema	c 936	16	2.0	3255	22	AAF55244	Nucleotide sequenc
c 864	16	2.0	2318	22	AA540855	cDNA encoding nove	c 937	16	2.0	3255	22	AAF55258	Nucleotide sequenc
c 865	16	2.0	2343	20	AAZ24811	Human secreted pro	c 938	16	2.0	3255	22	AAF55259	Nucleotide sequenc
c 866	16	2.0	2348	22	AAH46914	cDNA encoding huma	c 939	16	2.0	3255	22	AAF55260	Nucleotide sequenc
867	16	2.0	2350	22	AA198954	Human excretory re	c 940	16	2.0	3282	21	AA50241	Arabidopsis thalia
868	16	2.0	2350	22	AA163304	Human kidney relat	c 941	16	2.0	3282	22	AAI67268	GABApie cDNA seque
869	16	2.0	2370	19	AAV68069	Maize-optmised DN	c 942	16	2.0	3306	13	AAQ22202	A. chrysogenum pho
870	16	2.0	2403	17	AAAT13955	Maize-optmised VI	c 943	16	2.0	3306	13	AAQ23005	Phosphoglycerate k
871	16	2.0	2403	18	AAAT4004	Maize-optmised-B.	c 944	16	2.0	3306	14	AAQ48534	PGK. Acremonium c
872	16	2.0	2403	19	AAV68065	Maize-optmised DN	c 945	16	2.0	3309	21	AAF21886	Human breast and o
873	16	2.0	2403	19	AAV16177	Maize-optmised DN	c 946	16	2.0	3345	23	ABL12550	Drosophila melanog
874	16	2.0	2424	23	AA575668	DNA encoding novel	c 947	16	2.0	3362	22	ABA18396	Human nervous syst
c 875	16	2.0	2434	22	AAK36971	Human musculoskele	c 948	16	2.0	3391	22	AAK88698	Mannanase coding s
c 876	16	2.0	2438	19	AAV38687	Mus musculus SOCS1	c 949	16	2.0	3422	22	AAK03358	Human reproductive
877	16	2.0	2444	9	AAH81254	cDNA sequence in p	c 950	16	2.0	3422	22	AAK03360	Human immunoglobul
878	16	2.0	2444	17	AAAT13964	VIP3A(a) synthetic	951	16	2.0	3422	22	AA528909	Human immunoglobul
879	16	2.0	2444	18	AAAT74011	B. cereus VIP3A(a)	c 952	16	2.0	3422	22	AA528911	Human phosphodiester
880	16	2.0	2444	19	AAV16190	DNA for maize opti	c 953	16	2.0	3427	21	AAA09591	Human cDNA sequenc
881	16	2.0	2502	21	AAAG62679	Human NeuroBl gene	954	16	2.0	3434	22	AAH14440	Drosophila melanog
882	16	2.0	2505	21	AAK99819	Human secreted pro	955	16	2.0	3486	23	ABL26562	Drosophila melanog
883	16	2.0	2506	22	AAK07439	Human reproductive	c 956	16	2.0	3540	23	ABL02062	Drosophila melanog
884	16	2.0	2506	22	AAI62735	Human breast or ov	c 957	16	2.0	3551	23	ABL17756	Drosophila melanog
885	16	2.0	2527	23	ABL21520	Drosophila melanog	958	16	2.0	3559	22	AAI60356	Human polynucleoti

|||||
Db 361 CAGACCGGACAGGAGGCGGATGAGCGACACACACAAACAGACACACACGCC 420
QY 421 AGTCCAGGAGCCAGTAATGGAGAGGCCCAAAAGAGAACCCAGAGCTGAAAGTCGGG 480
Db 421 AGTCCAGGAGCCAGTAATGGAGAGGCCCAAAAGAGAACCCAGAGCTGAAAGTCGGG 480
QY 481 ATCTACACCTGGGACGACAGAGAGATCAGGATACAGCTGAGATCCCACTGCCG 540
Db 481 ATCTACACCTGGGACGACAGAGAGATCAGGATACAGCTGAGATCCCACTGCCG 540
QY 541 ACATGGAAGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTG 600
Db 541 ACATGGAAGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTG 600
QY 601 GGTTCGGGCTCAAGTGAAGATAATACCTAAAGAGGACACTGTAAATGCGCAAGCA 560
Db 601 GGTTCGGGCTCAAGTGAAGATAATACCTAAAGAGGACACTGTAAATGCGCAAGCA 560
QY 661 GGTGAAGAGCAACCAAGTTTAAATGAAGACAAGCTGAAACAACCAAGCTGGTTTTAT 720
Db 661 GGTGAAGAGCAACCAAGTTTAAATGAAGACAAGCTGAAACAACCAAGCTGGTTTTAT 720
QY 721 ATTAGATATTGACTTAACATATCTCAATAAGTTTTCAGCTTTTACCAAAAAA 780
Db 721 ATTAGATATTGACTTAACATATCTCAATAAGTTTTCAGCTTTTACCAAAAAA 780
QY 781 A 781
Db 781 A 781

RESULT 2
AA537109
ID AA537109 standard; cDNA; 396 BP.
XX
AC AA537109;
XX
DT 17-DEC-2001 (first entry)
XX
DE Novel human diagnostic and therapeutic gene #167.
XX
KW Human; cancer; breast; lung; colon; prostate; cytostatic; diagnostic; ss.
XX
OS Homo sapiens.
XX
PN WO200166753-A2.
XX
PD 13-SEP-2001.
XX
PF 09-MAR-2001; 2001WO-US07787.
XX
PR 09-MAR-2000; 2000US-0188609.
XX
(CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis WA, Garcia PD, Sudduth-Klinger J;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Kassam A, Lamson G;
PI Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
PI Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;
XX
DR WPI; 2001-530177/58.
XX
XX New polynucleotides and polypeptides, useful for diagnosis and
PT treatment of breast, lung and colon cancer -
XX
XX Claim 1; Page 638; 1193pp; English.
PS
XX The invention relates to new polynucleotides and polypeptides, useful for
CC diagnosis and treatment of breast, lung and colon cancer. The sequences
CC can be used in detecting differentially expressed genes correlated with a
CC cancerous state of a mammalian cell, comprising detecting at least one

CC differentially expressed gene product in a test sample derived from a
CC cell suspected of being cancerous. They can also be used to inhibit
CC tumour growth by modulating expression of a gene product. AA536943-
CC AA539338 represent novel human diagnostic and therapeutic coding
CC sequences of the invention.
XX
SQ Sequence 396 BP; 82 A; 97 C; 146 G; 71 T; 0 other;

Query Match 50.1%; Score 391; DB 22; Length 396;
Best Local Similarity 100.0%; Pred. No. 1.2e-172; Indels 0; Gaps 0;
Matches 391; Conservative 0; Mismatches 0;

QY 2 CGCGCGAGCTGTGAGCCCGGCACTCGGGTCCCTGAGGTCTCGATTCTTCTCCGCTACTG 61
Db 1 CGCGCGAGCTGTGAGCCCGGCACTCGGGTCCCTGAGGTCTCGATTCTTCTCCGCTACTG 60
QY 62 AGACACGGGGGTAGTCCACAGGCAGATCCAACTGGGAGTTGAAGTGTGAGTGAGAGTG 121
Db 61 AGACACGGGGGTAGTCCACAGGCAGATCCAACTGGGAGTTGAAGTGTGAGTGAGAGTG 120
QY 122 AAGAGGAACCAAGCAGCTTCCCGAGGGTTGTGTGTCAGTGACTCAGAGTGAGAAGGCC 181
Db 121 AAGAGGAACCAAGCAGCTTCCCGAGGGTTGTGTGTCAGTGACTCAGAGTGAGAAGGCC 180
QY 182 TCGAAGTCTGCTCCTCTCATGCGGTGCCACGCCCATGGACCTTCTGTCTCGTCACGG 241
Db 181 TCGAAGTCTGCTCCTCTCATGCGGTGCCACGCCCATGGACCTTCTGTCTCGTCACGG 240
QY 242 CCATAACTAGGAGGAAGGAGGCGGAGAGTGGAGGGCTCAGGCGAAGCTGGGGTGTCT 301
Db 241 CCATAACTAGGAGGAAGGAGGCGGAGAGTGGAGGGCTCAGGCGAAGCTGGGGTGTCT 300
QY 302 GTTGGGGTATCCGATCCCAAGACACCTGGAACCCCGACAGAGATTCTTGACTCTCCC 361
Db 301 GTTGGGGTATCCGATCCCAAGACACCTGGAACCCCGACAGAGATTCTTGACTCTCCC 360
QY 362 AGACGGGACCAAGAGAGGAGCGGATGACGG 392
Db 361 AGACGGGA/CAGGAGAGGAGCGGATGACGG 391

RESULT 3
AAH64751
ID AAH64751 standard; cDNA; 762 BP.
XX
AC AAH64751;
XX
DT 11-SEP-2001 (first entry)
XX
DE Human secreted protein cDNA, SEQ ID NO: 27.
XX
KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KW GENSET; ss.
XX
OS Homo sapiens.
XX
PN WO200142451-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-IB01938.
XX
PR 08-DEC-1999; 99US-0169629.
PR 06-MAR-2000; 2000US-0187470.
XX
XX (GEST) GENSET.
PA
XX Dumas Milne Edwards J, Bougueleret L, Jobert S;
PI
XX WPI; 2001-367870/38.
DR
DR P-PSDB; AAG89148.
XX
PT Full length GENSET human nucleic acids encoding potentially secreted

PT	proteins, useful in gene therapy and vaccination against a variety of	
PT	diseases, and for diagnosis of those diseases -	
XX		
PS	Claim 7; Page 586; 921pp; English.	
XX		
CC	The invention relates to full length GENSET human nucleic acids encoding	
CC	potentially secreted proteins. The nucleic acids and the polypeptides	
CC	they encode may be used in the prevention, treatment and diagnosis of	
CC	diseases associated with inappropriate GENSET gene expression. For	
CC	example, they be used to treat disorders associated with decreased	
CC	GENSET gene expression by rectifying mutations or deletions in a	
CC	patient's genome that affect the activity of GENSET or by supplementing	
CC	the patients own production of GENSET polypeptides. Conversely,	
CC	antisense nucleic acid molecules may be administered to down regulate	
CC	GENSET expression by binding with the cells' own genes and preventing	
CC	their expression. The sense and antisense nucleic acids may also be	
CC	used as DNA probes in diagnostic assays to detect and quantitate the	
CC	presence of similar nucleic acid sequences in samples, and hence to	
CC	determine which patients may be in need of restorative therapy.	
CC	The GENSET polypeptides may be used as antigens in the production of	
CC	antibodies and in assays to identify modulators (agonists and	
CC	antagonists) of GENSET polypeptide expression and activity. The	
CC	present sequence is a GENSET nucleic acid of the invention.	
XX		
SQ	Sequence 762 BP; 220 A; 192 C; 186 G; 164 T; 0 other;	
	Query Match 49.9%; Score 390; DB 22; Length 762;	
	Best Local Similarity 100.0%; Pred. No. 3.4e-172;	
	Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	392 GACACACACAAACACAGAACCCACAGCCAGTCCAGGAGCCAGTAATGGAGCCCCA 451	
DB	370 GACACACACAAACACAGAACCCACAGCCAGTCCAGGAGCCAGTAATGGAGCCCCA 429	
QY	452 AAAAGAAGAACCCAGCAGCTCAAAAGTCGGGATCCTACACCTGGCGCAGCAGACAAAGA 511	
DB	430 AAAAGAAGAACCCAGCAGCTCAAAAGTCGGGATCCTACACCTGGCGCAGCAGACAAAGA 489	
QY	512 TCAGGTATACAGTGCAGTCCAGTGCAGCAGATGGAGGTGATCTGCAAGAGTCGATCA 571	
DB	490 TCAGGTATACAGTGCAGTCCAGTGCAGCAGATGGAGGTGATCTGCAAGAGTCGATCA 549	
QY	572 GTCAAACACCGGGGATAAATCTGGATTTCGGTTCGGCGCTCAAGGTGAAGATAATACCTA 631	
DB	550 GTCAAACACCGGGGATAAATCTGGATTTCGGTTCGGCGCTCAAGGTGAAGATAATACCTA 609	
QY	632 AAGAGGAACACTGTAAATGTCAGAACGAGTGAAGAGCAACCAACAGTTTAATGAAGA 691	
DB	610 AAGAGGAACACTGTAAATGTCAGAACGAGTGAAGAGCAACCAACAGTTTAATGAAGA 669	
QY	692 CAAGCTGAACACACGCAAGCTGGTTTATATTAGATATTTGACTTAAACTATCTCAATAA 751	
DB	670 CAAGCTGAACACACGCAAGCTGGTTTATATTAGATATTTGACTTAAACTATCTCAATAA 729	
QY	752 AGTTTTCAGCTTTTCACCAAAAAA 781	
DB	730 AGTTTTCAGCTTTTCACCAAAAAA 759	
RESULT 4		
AAF68151		
ID	AAF68151 standard; cDNA; 399 BP.	
XX		
AC	AAF68151;	
XX		
DT	12-APR-2001 (first entry)	
XX		
DE	Human lung tumour protein related nucleotide sequence SEQ ID NO:69.	
XX		
KW	Human; lung cancer; lung tumour; lung tumour protein; gene therapy;	
KW	lung cancer antigen; lung tumour-specific antigen; diagnosis; vaccine;	
KW	cytostatic; antisense inhibition; ss.	
XX		

Db 190 GTCAACACCGGGATTAATCTGGATTTCGGGTCCTCAAGGTGAAGATAATACCTA 249
QY 632 AAGAGGAACACTGTAAATGCCAGAGCAGGTGAAGAGCAACCAACCAAGTTTAAATGAAGA 691
Db 250 AAGAGGAACACTGTAAATGCCAGAGCAGGTGAAGAGCAACCAACCAAGTTTAAATGAAGA 309
QY 692 CAAGCTGAAACACGCAAGCTGTTTATATTAGATATTGACTTAAACTATCTCAATAA 751
Db 310 CAAGCTGAAACACGCAAGCTGTTTATATTAGATATTGACTTAAACTATCTCAATAA 369
QY 752 AGTTTTCAGCTTTCACCAA 771
Db 370 AGTTTTCAGCTTTCACCAA 389
RESULT 5
AAF68851
ID AAF68851 standard; cDNA; 457 BP.
AC AAF68851;
XX
DT 12-APR-2001 (first entry)
XX
DE Human lung tumour protein related nucleotide sequence SEQ ID NO:790.
XX
KW Human; lung cancer; lung tumour; lung tumour protein; gene therapy;
KW lung cancer antigen; lung tumour-specific antigen; diagnosis; vaccine;
KW cytostatic; antisense inhibition; ss.
XX
OS Homo sapiens.
XX
PN WO200100828-A2.
XX
PD 04-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18061.
XX
PR 30-JUN-1999; 99US-0346492.
PR 15-OCT-1999; 99US-0419356.
PR 17-DEC-1999; 99US-0466867.
PR 30-DEC-1999; 99US-0476300.
PR 06-MAR-2000; 2000US-0519642.
PR 22-MAR-2000; 2000US-0533077.
PR 10-APR-2000; 2000US-0546259.
PR 27-APR-2000; 2000US-0560406.
PR 05-JUN-2000; 2000US-0589184.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Bangur CS, Lodes MJ, Fanger GR, Vedvick TS, Carter D;
PI Retter MW, Mannion J;
XX
XX WPI; 2001-071488/08.
XX
XX Lung tumor-associated proteins and the nucleic acids that encode them,
PT useful for preventing, diagnosing and treating lung cancer -
XX
XX Claim 4; Page 412; 436pp; English.
XX
XX The present invention describes immunogenic portions of lung tumour-
CC associated proteins (I) and the nucleic acids (NAS) that encode them.
CC (I) have cytostatic activity and can be used in gene therapy, antisense
CC inhibition and in vaccines. The NAS and the lung tumour-associated
CC proteins they encode may be used in the prevention, treatment and
CC diagnosis of diseases associated with their inappropriate expression,
CC especially lung cancers. For example, the NAS may be administered to
CC treat diseases by rectifying mutations or deletions in a patient's genome
CC that affect the activity of the protein by expressing inactive proteins
CC or to supplement the patients own production of (I). Additionally, the
CC NAS may be used to produce the lung-tumour associated protein, according
CC to standard recombinant DNA methodology. Conversely, antisense NA
CC molecules may be administered to down regulate protein expression by

CC binding with the cells own genes and preventing their expression. The NA
CC and complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar NA sequences in
CC samples, and hence which patients may be in need of treatment for lung
CC cancer. The (I) may be used as antigens in the production of antibodies
CC and in assays to identify modulators (agonists and antagonists) of the
CC expression and activity of the protein. AAF68083 to AAF68878 and
CC AAB76848 to AAB76878 represent human lung tumour protein related
CC nucleotide and protein sequences which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 457 BP; 146 A; 105 C; 121 G; 85 T; 0 other;
Query Match 47.1%; Score 368; DB 22; Length 457;
Best Local Similarity 100.0%; Pred. No. 6.5e-162;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 392 GACACACAAACACAGCAACACAGCCAGTCCCGAGAGCCCGAGTATGGAGAGCCCCA 451
Db 90 GACACACAAACACAG/ACCACACAGCCAGTCCCGAGAGCCCGAGTATGGAGAGCCCCA 149
QY 452 AAAAGAAGAACACAGCTGAAAGTCGGGATCCTACACCTGGCGAGCAGACAGAGAAGA 511
Db 150 AAAAGAAGAACACAGCTGAAAGTCGGGATCCTACACCTGGCGAGCAGACAGAGAAGA 209
QY 512 TCAGGATACAGTGAGATCCAGTCCGCGACATGGAAGTGATCTGCAAGAGCTGCATCA 571
Db 210 TCAGGATACAGTGAGATCCAGTCCGCGACATGGAAGTGATCTGCAAGAGCTGCATCA 269
QY 572 GTCAAAACACCGGGGATAAATCTGGATTTCGGGTCGGGTCAGGTGAAGATAATACCTA 631
Db 270 GTCAAAACACCGGGGATAAATCTGGATTTCGGGTCGGGTCAGGTGAAGATAATACCTA 329
QY 632 AAGAGGAACACTGTAAATGCCAGAACAGGTGAAGAGCAACCAACCAAGTTTAAATGAAGA 691
Db 330 AAGAGGAACACTGTAAATGCCAGAACAGGTGAAGAGCAACCAACCAAGTTTAAATGAAGA 389
QY 692 CAAGCTGAAACCAACGCAAGCTGTTTATATTAGATATTGACTTAAACTATCTCAATAA 751
Db 390 CAAGCTGAAACCAACGCAAGCTGTTTATATTAGATATTGACTTAAACTATCTCAATAA 449
QY 752 AGTTTTCG 759
Db 450 AGTTTTCG 457
RESULT 6
AAF68852
ID AAF68852 standard; cDNA; 461 BP.
XX
AC AAF68852;
XX
DT 12-APR-2001 (first entry)
XX
DE Human lung tumour protein related nucleotide sequence SEQ ID NO:792.
XX
KW Human; lung cancer; lung tumour; lung tumour protein; gene therapy;
KW lung cancer antigen; lung tumour-specific antigen; diagnosis; vaccine;
KW cytostatic; antisense inhibition; ss.
XX
OS Homo sapiens.
XX
PN WO200100828-A2.
XX
PD 04-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18061.
XX
PR 30-JUN-1999; 99US-0346492.
PR 15-OCT-1999; 99US-0419356.
PR 17-DEC-1999; 99US-0466867.
PR 30-DEC-1999; 99US-0476300.
PR 06-MAR-2000; 2000US-0519642.
PR

```

PR 22-MAR-2000; 2000US-0533077.
PR 10-APR-2000; 2000US-0546259.
PR 27-APR-2000; 2000US-0560406.
PR 05-JUN-2000; 2000US-0589184.
XX
XX (CORI-) CORIXA CORP.
XX
XX Wang T, Bangur CS, Lodes MJ, Fanger GR, Vedvick TS, Carter D;
XX Retter MW, Mannion J;
XX
XX WPI: 2001-071488/08.
XX
XX Lung tumor-associated proteins and the nucleic acids that encode them,
XX useful for preventing, diagnosing and treating lung cancer -
XX
XX Claim 4; Page 413; 436pp; English.
XX
XX The present invention describes immunogenic portions of lung tumour-
XX associated proteins (I) and the nucleic acids (NAs) that encode them.
XX (I) have cytostatic activity and can be used in gene therapy, antisense
XX inhibition and in vaccines. The NAs and the lung tumour-associated
XX proteins they encode may be used in the prevention, treatment and
XX diagnosis of diseases associated with their inappropriate expression,
XX especially lung cancers. For example, the NAs may be administered to
XX treat diseases by rectifying mutations or deletions in a patient's genome
XX that affect the activity of the protein by expressing inactive proteins
XX or to supplement the patients own production of (I). Additionally, the
XX NAs may be used to produce the lung-tumour associated protein, according
XX to standard recombinant DNA methodology. Conversely, antisense NA
XX molecules may be administered to down regulate protein expression by
XX binding with the cells own genes and preventing their expression. The NA
XX and complementary sequences may also be used as DNA probes in diagnostic
XX assays to detect and quantitate the presence of similar NA sequences in
XX samples, and hence which patients may be in need of treatment for lung
XX cancer. The (I) may be used as antigens in the production of antibodies
XX and in assays to identify modulators (agonists and antagonists) of the
XX expression and activity of the protein. AAF68083 to AAF68878 and
XX AAB76848 to AAB76878 represent human lung tumour protein related
XX nucleotide and protein sequences which are used in the exemplification
XX of the present invention.
XX
XX Sequence 461 BP; 150 A; 104 C; 123 G; 84 T; 0 other;
XX
XX Query Match 29.7%; Score 232; DB 22; Length 461;
XX Best Local Similarity 100.0%; Pred. No. 2e-98;
XX Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 535 TCGCGGACATGGAAGGTGATCTGCAAGAGCTGCATCAAGTCAAAACACGGGGATAAATCTG 594
DB 230 TCGCGGACATGGAAGGTGATCTGCAAGAGCTGCATCAAGTCAAAACACGGGGATAAATCTG 289
QY 595 GATTTGGTTCGGCGTCAAGGTGAAGATAATACCTAAAGAGGAACTGTAAATGCCA 654
DB 290 GATTTGGTTCGGCGTCAAGGTGAAGATAATACCTAAAGAGGAACTGTAAATGCCA 349
QY 655 GAAGCAGGTGAAGAGCAACCAACAACTTTAATGAAGACAAGCTGAAACACCGCAAGCTGG 714
DB 350 GAAGCAGGTGAAGAGCAACCAACAACTTTAATGAAGACAAGCTGAAACACCGCAAGCTGG 409
QY 715 TTTTATATTAGATATTGACTTAACCTATCTCAATAAAGTTTTCAGCTTTC 766
DB 410 TTTTATATTAGATATTGACTTAACCTATCTCAATAAAGTTTTCAGCTTTC 461
XX
RESULT 7
ID AAF68152
XX AAF68152 standard; cDNA; 479 BP.
XX
XX AAF68152;
XX
XX 12-APR-2001 (first entry)
XX
XX Human lung tumour protein related nucleotide sequence SEQ ID NO:70.

```

```

XX
XX Human; lung cancer; lung tumour; lung tumour protein; gene therapy;
XX lung cancer antigen; lung tumour-specific antigen; diagnosis; vaccine;
XX cytostatic; antisense inhibition; ss.
XX
XX Homo sapiens.
XX
XX WO200100828-A2.
XX
XX 04-JAN-2001.
XX
XX 30-JUN-2000; 2000WO-US18061.
XX
XX 30-JUN-1999; 99US-0346492.
XX 15-OCT-1999; 99US-0419356.
XX 17-DEC-1999; 99US-0466867.
XX 30-DEC-1999; 99US-0476300.
XX 06-MAR-2000; 2000US-0519642.
XX 22-MAR-2000; 2000US-0533077.
XX 10-APR-2000; 2000US-0546259.
XX 27-APR-2000; 2000US-0560406.
XX 05-JUN-2000; 2000US-0589184.
XX
XX (CORI-) CORIXA CORP.
XX
XX Wang T, Bangur CS, Lodes MJ, Fanger GR, Vedvick TS, Carter D;
XX Retter MW, Mannion J;
XX
XX WPI: 2001-071488/08.
XX
XX Lung tumor-associated proteins and the nucleic acids that encode them,
XX useful for preventing, diagnosing and treating lung cancer -
XX
XX Claim 4; Page 174; 436pp; English.
XX
XX The present invention describes immunogenic portions of lung tumour-
XX associated proteins (I) and the nucleic acids (NAs) that encode them.
XX (I) have cytostatic activity and can be used in gene therapy, antisense
XX inhibition and in vaccines. The NAs and the lung tumour-associated
XX proteins they encode may be used in the prevention, treatment and
XX diagnosis of diseases associated with their inappropriate expression,
XX especially lung cancers. For example, the NAs may be administered to
XX treat diseases by rectifying mutations or deletions in a patient's genome
XX that affect the activity of the protein by expressing inactive proteins
XX or to supplement the patients own production of (I). Additionally, the
XX NAs may be used to produce the lung-tumour associated protein, according
XX to standard recombinant DNA methodology. Conversely, antisense NA
XX molecules may be administered to down regulate protein expression by
XX binding with the cells own genes and preventing their expression. The NA
XX and complementary sequences may also be used as DNA probes in diagnostic
XX assays to detect and quantitate the presence of similar NA sequences in
XX samples, and hence which patients may be in need of treatment for lung
XX cancer. The (I) may be used as antigens in the production of antibodies
XX and in assays to identify modulators (agonists and antagonists) of the
XX expression and activity of the protein. AAF68083 to AAF68878 and
XX AAB76848 to AAB76878 represent human lung tumour protein related
XX nucleotide and protein sequences which are used in the exemplification
XX of the present invention.
XX
XX Sequence 479 BP; 163 A; 107 C; 125 G; 84 T; 0 other;
XX
XX Query Match 24.5%; Score 191; DB 22; Length 479;
XX Best Local Similarity 100.0%; Pred. No. 2.8e-79;
XX Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 535 TCGCGGACATGGAAGGTGATCTGCAAGAGCTGCATCAAGTCAAAACACGGGGATAAATCTG 594
DB 232 TCGCGGACATGGAAGGTGATCTGCAAGAGCTGCATCAAGTCAAAACACGGGGATAAATCTG 291
QY 595 GATTTGGTTCGGCGTCAAGGTGAAGATAATACCTAAAGAGGAACTGTAAATGCCA 654
DB 292 GATTTGGTTCGGCGTCAAGGTGAAGATAATACCTAAAGAGGAACTGTAAATGCCA 351

```

QY 655 GAAGCAGGTGAAGGACACACCAAGTTTAAATGAAGACAAAGCTGAACCAACCAAGCTGG 714
Db 352 GAAGCAGGTGAAGGACACACCAAGTTTAAATGAAGACAAAGCTGAACCAACCAAGCTGG 411
QY 715 TTTTATATTAG 725
Db 412 TTTTATATTAG 422

RESULT 8
AA442613/c
ID AAA42613 standard; cDNA; 201 BP.
XX
AC AAA42613;
XX
DT 21-AUG-2000 (first entry)
XX
DE Human secreted expressed sequence tag SEQ ID NO:1353.
XX
KW Human; mouse; xenopus; rat; secreted expressed sequence tag; sEST;
KW expressed sequence tag; EST; probe; chemotactic; proliferative;
KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
KW thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;
KW antiviral; antidiabetic; antiasthmatic; vulnary; antiparkinsonian;
KW antiulcer; osteopathic; neuroprotective; nootropic; antipsoriatic;
KW cerebroprotective; anticonvulsant; antidepressant; gene therapy;
KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;
KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
KW central nervous system disorder; Alzheimer's disease; stroke;
KW parkinson's disease; Huntington's disease; coagulation disorder;
KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;
KW tumour; infection; depression; psoriasis; ss.
XX
OS Homo sapiens.
XX
PN WO200021990-A1.
XX
PD 20-APR-2000.
XX
PF 15-OCT-1999; 99WO-US24205.
XX
PR 15-OCT-1998; 98US-0104435.
XX
PA (GENY) GENETICS INST INC.
XX
XX Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;
PI Merberg D, Treacy M;
XX
DR WPI; 2000-317937/27.
XX
PT Isolated polynucleotides, and encoded proteins, comprising secreted
PT expressed sequence tags (sESTs), useful for treating various disorders
PT such as autoimmune, infectious, and central nervous system disorders -
XX
XX Claim 1; Page 447; 618pp; English.
XX
XX AAA41261 to AAA43419 represent specifically claimed secreted expressed
CC sequence tags (sESTs), isolated from human, mouse, xenopus and rat
CC tissue sources. The sESTs can have a range of activities depending on
CC the tissues they were isolated from. The activities include:
CC chemotactic; proliferative; immunomodulatory; haematopoietic;
CC chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;
CC cytostatic; antibacterial; antifungal; antiviral; antidiabetic;
CC antiasthmatic; vulnary; antiulcer; osteopathic; neuroprotective;
CC nootropic; antiparkinsonian; antipsoriatic; cerebroprotective;
CC anticonvulsant; and antidepressant. The sESTs can be used for gene
CC therapy and in vaccines. The sESTs are useful as probes for the
CC identification and isolation of full-length cDNAs and genomic DNA
CC molecules which correspond to the sESTs. Proteins encoded by the sESTs
CC are useful in assays for determining biological activity and raising
CC antibodies. They may be useful for treatment of autoimmune disorders
CC (multiple sclerosis, insulin dependent diabetes), allergic conditions

CC (asthma), myeloid or lymphoid cell deficiencies, wounds, ulcers,
CC osteoporosis, osteoarthritis, central nervous system disorders
CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
CC disease), tumours, bacterial, fungal or viral infections, depression and
CC psoriasis. AAA43420 to AAA43425 represent linker variants which are given
CC in the exemplification of the present invention.
XX
SQ Sequence 201 BP; 33 A; 56 C; 42 G; 70 T; 0 other;
Query Match 22.3%; Score 174; DB 21; Length 201;
Best Local Similarity 100.0%; Pred. No. 2.5e-71;
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 538 GGCACATGGAAGGTGATCTGCAAGAGCTGCATCAAGACACCGGGGATAAATCTGGAT 597
Db 194 GGCACATGGAAGGTGATCTGCAAGAGCTGCATCAAGACACCGGGGATAAATCTGGAT 135
QY 598 TTGGGTTCGGCGTCAAGGTGAAGATAATACCTAAAGAGGAACACTGTAAAATGCCAGAA 657
Db 134 TTGGGTTCGGCGTCAAGGTGAAGATAATACCTAAAGAGGAACACTGTAAAATGCCAGAA 75
QY 658 GCAGGTGAAGAGCAACCAACCAAGTTTAAATGAAGACAAAGCTGAACCAACCAAGC 711
Db 74 GCAGGTGAAGAGCAACCAACCAAGTTTAAATGAAGACAAAGCTGAACCAACCAAGC 21

RESULT 9
AAL22641
ID AAL22641 standard; cDNA; 822 BP.
XX
AC AAL22641;
XX
DT 07-DEC-2001 (first entry)
XX
DE Human breast cancer expressed polynucleotide 15098.
XX
KW Human; breast cancer; cell marker; cytostatic; ss.
XX
OS Homo sapiens.
XX
PN WO200151628-A2.
XX
PD 19-JUL-2001.
XX
PF 10-JAN-2001; 2001WO-US00798.
XX
PR 14-JAN-2000; 2000US-0176077.
PR 14-MAR-2000; 2000US-0189167.
PR 24-MAR-2000; 2000US-0192099.
PR 29-MAR-2000; 2000US-0193480.
PR 15-MAY-2000; 2000US-0205230.
PR 09-JUN-2000; 2000US-0211315.
PR 25-JUL-2000; 2000US-0220534.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX WPI; 2001-451856/48.
XX
XX New peptide useful as a marker for the diagnosis of breast cancer -
PT
XX Claim 1; Page 2729-2730; 3695pp; English.
XX
XX The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic

QY 582 GCGGATAAATCTGGATTTCGGTTCCGGCTCAAGGTGAAGATAATACCTAAAGAGGAACA 641
Db 111 GGGGATAAATCTGGATTTCGGTTCCGGCTCAAGGTGAAGATAATACCTAAAGAGGAACA 170
QY 642 CTCTAAATGCCAGAGCAGGTGAAGACCAACCAAGTTTAAATGAAGACAAAGCTGAAA 701
Db 171 CTCTAAATGCCAGAGCAGGTGAAGACCAACCAAGTTTAAATGAAGACAAAGCTGAAA 230
QY 702 CAACGCAAGCTGTTTTATATTAGATATTGACTTAAACTATCTCAATAAAGTTTTCGAG 761
Db 231 CAACGCAAGCTGTTTTATATTAGATATTGACTTAAACTATCTCAATAAAGTTTTCGAG 290
QY 762 CTTTCACCA 770
Db 291 CTTTCACCA 299
RESULT 12
AAC10552
ID AAC10552 standard; cDNA; 190 BP.
XX
AC AAC10552;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 14627.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX Homo sapiens.
XX
XX EPI033401-A2.
XX
XX 06-SEP-2000.
XX
XX 21-FEB-2000; 2000EP-0200610.
XX
XX 26-FEB-1999; 99US-0122487.
XX
XX (GEST) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI; 2000-500381/45.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1; SEQ ID 14627; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 190 BP; 52 A; 51 C; 58 G; 24 T; 5 other;
Query Match 11.5%; Score 90; DB 21; Length 190;
Best Local Similarity 100.0%; Pred. No. 4e-32;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 GACACACACAACACAGCAAGCCACAGCCAGTCCAGGAGCCAGTAATGAGAGCCCA 451
Db 83 GACACACACAACACAGCAAGCCACAGCCAGTCCAGGAGCCAGTAATGAGAGCCCA 142
QY 452 AAAAGAGAACACAGCAGCTGAAAGTCGGGA 481
Db 143 AAAAGAGAACACAGCAGCTGAAAGTCGGGA 172
RESULT 13
AAL13774
ID AAL13774 standard; cDNA; 618 BP.
XX
AC AAL13774;
XX
DT 07-DEC-2001 (first entry)
XX
DE Human breast cancer expressed polynucleotide 6231.
XX
KW Human; breast cancer; cell marker; cytostatic; ss.
XX
XX Homo sapiens.
XX
XX WO200151628-A2.
XX
XX 19-JUL-2001.
XX
XX 10-JAN-2001; 2001WO-US00798.
XX
XX 14-JAN-2000; 2000US-0176077.
XX
XX 14-MAR-2000; 2000US-0189167.
XX
XX 24-MAR-2000; 2000US-0192099.
XX
XX 29-MAR-2000; 2000US-0193480.
XX
XX 15-MAY-2000; 2000US-0205230.
XX
XX 09-JUN-2000; 2000US-0211315.
XX
XX 25-JUL-2000; 2000US-0220534.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX
XX WPI; 2001-451856/48.
XX
PT New peptide useful as a marker for the diagnosis of breast cancer -
XX
XX Claim 1; Page 1118; 3695pp; English.
XX
CC The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity.
XX
SQ Sequence 618 BP; 169 A; 141 C; 144 G; 119 T; 45 other;
Query Match 8.6%; Score 67; DB 22; Length 618;
Best Local Similarity 100.0%; Pred. No. 2.1e-21;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCGGAGCTGTGAGCGCGGAGCTCGGTCCTCGAGTCTGGATTCTTCTCCGCTACT 60
Db 65 GCGCGGAGCTGTGAGCGCGGAGCTCGGTCCTCGAGTCTGGATTCTTCTCCGCTACT 124
QY 61 GAGACAC 67
Db 125 GAGACAC 131

```

RESULT 14
AA45936
ID   AAA45936 standard; DNA; 1092 BP.
XX
XX
AC   AAA45936;
XX
XX
DT   23-AUG 2000 (first entry)
XX
DE   Human metastatic marker gene sequence SEQ ID NO:5.
XX
KW   Human; breast cancer; colon cancer; metastatic marker; metastasis;
KW   tumour; cytostatic; gene therapy; diagnosis; detection; identification;
KW   screening; ds.
XX
XX
OS   Homo sapiens.
XX
XX
PN   WO200022130-A2.
XX
XX
PD   20-APR-2000.
XX
XX
PF   14-OCT-1999; 99WO-US24222.
XX
XX
PR   15-OCT-1998; 98US-0104351.
XX
PR   13-OCT-1999; 99US-0104351.
XX
XX
PA   (CHTR ) CHIRON CORP.
XX
PI   Giese K;
XX
XX
DR   WPI; 2000-317985/27.
XX
XX
PT   Novel human proteins encoded by metastatic marker genes which are
PT   differentially expressed between metastatic cancer cells esp., of
PT   breast and colon, and non-metastatic cancer cells, useful in
PT   diagnostics -
XX
XX
PS   Claim 3; Page 60-61; 89pp; English.
XX
XX
CC   AAA45932 to AAA46016 represent human metastatic marker gene sequences.
CC   The present invention describes: (1) isolated and purified human protein
CC   sequences encoded by AAA45932 to AAA45994; and (2) methods for detecting
CC   metastatic tumour cells and/or determining metastatic potential in a
CC   tissue sample. The metastatic marker polynucleotide sequence can be
CC   used as diagnostic agents for detecting metastatic tumour cells,
CC   determining metastatic potential in a tissue sample, propensity for the
CC   metastatic spread of a breast tumour preferentially to bone or lung, or
CC   the metastatic spread of a colon tumour. They can be used as markers of
CC   metastasis and are useful for rationally prescribing the course of
CC   therapy for breast or colon cancer patients. Detection of metastatic
CC   marker gene expression is useful for e.g. identifying metastasis or for
CC   determining metastatic potential in a tissue sample, preferably a
CC   tumour. Appropriate treatment regimens can then be designed for patients
CC   who are at risk for developing metastatic cancers in other organs of the
CC   body. The marker protein is also used for screening drugs which have a
CC   therapeutic anti-metastatic effect. Antibodies which specifically bind
CC   to the metastatic markers can also be used to alter metastatic marker
CC   gene expression, e.g. for treating cancers.
XX
XX
SQ   Sequence 1092 BP; 142 A; 305 C; 125 G; 272 T; 248 other;
XX
XX
Query Match 5.6%; Score 44; DB 21; Length 1092;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 577 ACACCGGGGATAAATCTGGATTGGGTTCCGGCGTCGAAGGTGAA 620
|||||
DB 79 ACACCGGGGATAAATCTGGATTGGGTTCCGGCGTCGAAGGTGAA 122
|||||

RESULT 15
AAC59112/c
ID   AAC59112 standard; cDNA; 2182 BP.
XX
XX

```

```

AAC59112;
XX
XX
DT   02-FEB-2001 (first entry)
XX
DE   Human secreted protein coding sequence SEQ ID NO: 15.
XX
XX
KW   Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW   antiallergic; hepatotropic; antidiabetic; antinflammatory; antitumor;
KW   vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW   cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; human; secreted protein; ss.
XX
XX
OS   Homo sapiens.
XX
XX
PN   WO200055177-A2.
XX
XX
PD   21-SEP-2000.
XX
XX
PF   09-MAR-2000; 2000WO-US06058.
XX
XX
PR   12-MAR-1999; 99US-0124145.
XX
PR   03-DEC-1999; 99US-0168654.
XX
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI   Rosen CA, Ruben SM, Komatsoulis G;
XX
XX
DR   WPI; 2000-638177/61.
XX
XX
DR   P-PSDB; AAB28016.
XX
XX
PT   Novel nucleic acids encoding 49 human secreted proteins useful for
PT   treating cancers, hyperproliferative disorders, inflammatory disorders,
PT   neurological disorders and cardiovascular disorders -
XX
XX
PS   Claim 1; Page 319-320; 389pp; English.
XX
XX
CC   The invention relates to the isolation of genes AAA59108-A59156 encoding
CC   49 human secreted proteins AAB28012-B28060. The genes can be used to
CC   generate fusion proteins by linking to the gene for the human
CC   immunoglobulin G Fc portion (SEQIDP1) for increasing the stability of
CC   the fusion protein as compared to the human protein only. The genes and
CC   proteins are useful for preventing, ameliorating or treating medical
CC   conditions, e.g. by protein or gene therapy. The genes are isolated
CC   from a range of human tissues disclosed in the specification. The
CC   nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC   the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC   and ovarian cancer, and other cancers of the adrenal gland, bone, bone
CC   marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC   (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC   disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC   colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d)
CC   wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC   epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC   and parasitic infections.
XX
XX
SQ   Sequence 2182 BP; 502 A; 639 C; 579 G; 452 T; 0 other;
XX
XX
Query Match 3.7%; Score 29; DB 21; Length 2182;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 ATCCAACCTGGGAGTTGAAGTGTGAGTGAG 117
|||||
DB 156 ATCCAACCTGGGAGTTGAAGTGTGAGTGAG 128
|||||

RESULT 16
AAS69484
ID   AAS69484 standard; cDNA; 503 BP.
XX
XX
AC   AAS69484;
XX
XX

```

DT 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #5288.
DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
KW Homo sapiens.
XX WO200175067-A2.
PN 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
PA Drmanac RT, Liu C, Tang YT;
PI WPI: 2001-639362/73.
DR P-PSDB; ABG05297.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX Claim 1; SEQ ID NO 5288; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 503 BP; 158 A; 98 C; 121 G; 105 T; 21 other; .
Query Match 3.5%; Score 27; DB 23; Length 503;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 679 GTTTAAATGAAGACAGCTGAACAC 705
Db 390 GTTTAAATGAAGACAGCTGAACAC 416
RESULT 17
AAI60530
ID AAI60530 standard; cDNA: 532 BP.
XX
AC AAI60530;
XX
DT 22-OCT-2001 (first entry)
XX

DE Human polynucleotide SEQ ID NO 4519.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX WO200153312-A1.
PN 26-JUL-2001.
XX 26-DEC-2000; 2000WO-US34263.
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
DR P-PSDB; AAM41374.
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
PS Claim 1; SEQ ID NO 4519; 10078pp; English.
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 532 BP; 154 A; 108 C; 150 G; 120 T; 0 other;
Query Match 3.5%; Score 27; DB 22; Length 532;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 679 GTTTAAATGAAGACAGCTGAACAC 705
Db 433 GTTTAAATGAAGACAGCTGAACAC 459
RESULT 18
AAI58744
ID AAI58744 standard; cDNA: 659 BP.
XX
AC AAI58744;

XX	22-OCT-2001 (first entry)	AA569486 standard; cDNA; 661 BP.
XX	Human polynucleotide SEQ ID NO 947.	AA569486;
XX	Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia; ss.	13-FEB-2002 (first entry)
XX	Homo sapiens.	DNA encoding novel human diagnostic protein #5290.
XX	WO200153312-A1.	Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX	26-JUL-2001.	Homo sapiens.
XX	26-DEC-2000; 2000WO-US34263.	WO200175067-A2.
XX	21-JAN-2000; 2000US-0488725.	11-OCT-2001.
XX	25-APR-2000; 2000US-0552317.	30-MAR-2001; 2001WO-US08631.
XX	09-JUL-2000; 2000US-0598042.	31-MAR-2000; 2000US-0540217.
XX	19-JUL-2000; 2000US-0620312.	23-AUG-2000; 2000US-0649167.
XX	03-AUG-2000; 2000US-0653450.	(HYSE-) HYSEQ INC.
XX	14-SEP-2000; 2000US-0662191.	Drmanac RT, Liu C, Tang YT;
XX	19-OCT-2000; 2000US-0693036.	WPI; 2001-639362/73.
XX	29-NOV-2000; 2000US-0727344.	P-PSDB; ABG05299.
XX	(HYSE-) HYSEQ INC.	New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity
XX	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D; Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J; Zhao QA, Zhou P, Goodrich R, Drmanac RT;	Claim 1; SEQ ID No 5290; 103pp; English.
XX	WPI; 2001-442253/47.	The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention.
XX	P-PSDB; AAM39588.	Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX	Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries	Sequence 661 BP; 212 A; 129 C; 158 G; 162 T; 0 other;
XX	Claim 1; SEQ ID NO 947; 10078pp; English.	Query Match 3.5%; Score 27; DB 23; Length 661;
XX	The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAM38642-AAM42213) with neurotropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders.	Best Local Similarity 100.0%; Pred. No. 0.01;
XX	Note: The sequence data for this patent did not form part of the printed specification.	Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX	Sequence 659 BP; 197 A; 137 C; 173 G; 152 T; 0 other;	Qy 679 GTTTAAATGAAGACAAGCTGAACAAC 705
XX	Query Match 3.5%; Score 27; DB 22; Length 659;	Db 563 GTTTAAATGAAGACAAGCTGAACAAC 589
XX	Best Local Similarity 100.0%; Pred. No. 0.01;	RESULT 19
XX	Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	AA569486
XX	Qy 679 GTTTAAATGAAGACAAGCTGAACAAC 705 	AA538132
XX	Db 555 GTTTAAATGAAGACAAGCTGAACAAC 581 	XX

```

AC AAS38132;
XX
DT 17-DEC-2001 (first entry)
XX
XX Novel human diagnostic and therapeutic gene #1190.
XX
XX Human; cancer; breast; lung; colon; prostate; cytostatic; diagnostic; ss.
XX
XX Homo sapiens.
XX
XX WO200166753-A2.
XX
XX 13-SEP-2001.
XX
XX 09-MAR-2001; 2001WO-US07787.
XX
XX 09-MAR-2000; 2000US-0188609.
XX
XX (CHIR ) CHIRON CORP.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
PI Reinhard C, Randazzo F, Kennedy CC, Pot D, Kassam A, Lamson G;
PI Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
PI Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;
XX
XX WPI; 2001-530177/58.
XX
XX New polynucleotides and polypeptides, useful for diagnosis and
PT treatment of breast, lung and colon cancer -
XX
XX Claim 1; Page 895; 1193pp; English.
XX
XX The invention relates to new polynucleotides and polypeptides, useful for
CC diagnosis and treatment of breast, lung and colon cancer. The sequences
CC can be used in detecting differentially expressed genes correlated with a
CC cancerous state of a mammalian cell, comprising detecting at least one
CC differentially expressed gene product in a test sample derived from a
CC cell suspected of being cancerous. They can also be used to inhibit
CC tumour growth by modulating expression of a gene product. AAS36943-
CC AAS39338 represent novel human diagnostic and therapeutic coding
CC sequences of the invention.
XX
XX Sequence 391 BP; 82 A; 109 C; 117 G; 83 T; 0 other;
SQ
Query Match 3.1%; Score 24; DB 22; Length 391;
Best Local Similarity 100.0%; Pred. No. 0.26;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 169 AGTGAGAGGCCCTCGAAGTCGTC 192
DB 169 AGTGAGAGGCCCTCGAAGTCGTC 192
|||||
RESULT 21
AAS24637
ID AAS24637 standard; cDNA; 214 BP.
XX
XX AAS24637;
AC
XX
XX 07-NOV-2001 (first entry)
DT
XX
XX Human ovarian PCR-subtracted cDNA library clone #818.
DE
XX
XX Immunogenic protein; cancer; ovarian tumour; T-cell stimulation; ss;
KW gene therapy; cytostatic; T-cell expansion; nucleic acid hybridisation;
KW primer; probe.
XX
XX Homo sapiens.
OS
XX
XX WO200157207-A2.
PN
XX
XX 09-AUG-2001.
PD
AAH83260
ID AAH83260 standard; cDNA; 320 BP.
XX
XX AAH83260;
AC
XX
XX 25-SEP-2001 (first entry)
DT
XX
XX Human ovarian tumour associated polynucleotide sequence SEQ ID NO:884.
DE
XX
XX Human; ovarian tumour; ovarian cancer; diagnosis; gene therapy;
KW immunogenic; vaccine; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200151513-A2.
PN
XX
XX 19-JUL-2001.
PD
XX
XX 16-JAN-2001; 2001WO-US01575.
XX
XX 14-JAN-2000; 2000US-0176722.
XX
XX (CORI-) CORIXA CORP.
PA
XX
XX Algate PA;
PI
XX
XX WPI; 2001-425866/45.
DR
XX
XX
XX

```

05-FEB-2001; 2001WO-US03733.

04-FEB-2000; 2000US-0180403.

28-MAR-2000; 2000US-0192745.

(CORI-) CORIXA CORP.

Algate PA, Mannion J;

WPI; 2001-488879/53.

New polynucleotides encoding ovarian tumour proteins, useful for treating ovarian cancer, and as probes, primers, and markers of cancer progression -

Example 1; page 241-242; 378pp; English.

The invention comprises compositions used for the therapy and diagnosis of ovarian cancer. The compositions comprise one or more ovarian tumour proteins, their associated polynucleotides, or immunogenic portions of the proteins. The ovarian tumour polynucleotides and polypeptides are useful for stimulating and/or expanding T cells specific for a tumour protein. They are also useful for inhibiting the development of cancer in a patient with an ovarian tumour DNA or protein by incubating isolated T-cells allowing them to proliferate, and administering to the patient. The sequences can be used as markers for cancer, for example, to monitor ovarian cancer progression. Probes and primers are useful in nucleic acid hybridisation, in detecting the presence of complementary sequences in a given sample, for preparing mutant species and for preparing other genetic constructions. Sequences AAS23820-AAS25231 and AAS25328-AAS25549 represent human ovarian tumour protein cDNA clones.

Sequence 214 BP; 74 A; 34 C; 55 G; 51 T; 0 other;

Query Match 2.9%; Score 23; DB 22; Length 214;

Best Local Similarity 100.0%; Pred. No. 0.76;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

545 TAAATGCCAGAGCAGGTGAAG 667

|||||

122 TAAATGCCAGAGCAGGTGAAG 144

RESULT 22

AAH83260

ID AAH83260 standard; cDNA; 320 BP.

AAH83260;

25-SEP-2001 (first entry)

Human ovarian tumour associated polynucleotide sequence SEQ ID NO:884.

Human; ovarian tumour; ovarian cancer; diagnosis; gene therapy;

immunogenic; vaccine; ss.

Homo sapiens.

WO200151513-A2.

19-JUL-2001.

16-JAN-2001; 2001WO-US01575.

14-JAN-2000; 2000US-0176722.

(CORI-) CORIXA CORP.

Algate PA;

WPI; 2001-425866/45.

PT Novel ovarian tumor proteins, and nucleic acids encoding them, used to
PT treat and diagnose cancers, particularly ovarian cancer .
PS Claim 5; Page 225; 338pp; English.
XX
CC AAH82377 to AAH83878 represent human ovarian tumour-associated
CC polynucleotide sequences which encode ovarian tumour proteins. The
CC ovarian tumour protein and polynucleotide sequences have cytostatic
CC activity, and can be used in gene therapy and vaccine production. The
CC ovarian tumour proteins and polynucleotides can be used to inhibit
CC the development of cancer, particularly ovarian cancer. They can also
CC be used to diagnose the onset and progression of cancer.
XX
XX Sequence 320 BP; 96 A; 65 C; 85 G; 74 T; 0 other;
SQ
Query Match 2.9%; Score 23; DB 22; Length 320;
Best Local Similarity 100.0%; Pred. No. 0.75; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0;
Qy 645 TAAATGCCAGAGCAGGTGAAG 667
Db 122 TAAATGCCAGAGCAGGTGAAG 144
RESULT 23
AAF59637
ID AAF59637 standard; cDNA: 580 BP.
AC AAF59637;
XX
XX 24-APR-2001 (first entry)
DT
XX Human cell cycle and proliferation protein CCYPR-48 cDNA, SEQ ID NO:102.
DE
XX Cell cycle and proliferation protein; CCYPR: human; agonist;
KW antagonist; gene therapy; detection; gene therapy;
KW transgenic animal disease model; immune disorder;
KW developmental disorder; cell signalling disorder;
KW cell proliferative disorder; cancer; tumour; anaemia; epilepsy;
KW arteriosclerosis; asthma; allergy; diabetes mellitus;
KW menstrual cycle disorder; bacterial infection; ss.
XX Homo sapiens.
OS
XX WO200107471-A2.
PN
XX 01-FEB-2001.
PD
XX 21-JUL-2000; 2000WO-US19948.
PF
XX 21-JUL-1999; 99US-0145075.
PR
XX 08-SEP-1999; 99US-0153129.
PR
XX 10-NOV-1999; 99US-0164647.
XX
XX (INCY-) INCYTE GENOMICS INC.
PA
XX Hillman JL, Lal P, Tang YT, Yue H, Au-Young J, Bandman O;
PI Azimzai Y, Yang J, Lu DAM, Baughn MR, Patterson C, Shah P;
XX
XX WPI: 2001-112727/12.
DR
XX P-PSDB; AAB60500.
XX
XX Human cell cycle and proliferation proteins and polynucleotides are
PT used to treat, diagnose and prevent immune, developmental and cell
PT signalling disorders and cell proliferative disorders including cancer .
XX
XX Claim 5; Page 200; 205pp; English.
PS
XX Sequences AAF59590-AAF59643 represent cDNAs encoding 54 human
CC cell cycle and proliferation proteins (CCYPR), AAB60453-AAB60506.
CC CCYPR and agonists of CCYPR are used to treat diseases or conditions
CC associated with decreased expression of functional CCYPR, while CCYPR
CC antagonists are used to treat diseases or conditions associated with

CC overexpression of functional CCYPR. Monoclonal or polyclonal antibodies
CC to CCYPR may be used in enzyme-linked immunosorbent assays (ELISA) or
CC radioimmunoassays to detect CCYPR. CCYPR itself may be used to detect
CC compounds e.g., antibodies, oligonucleotides and proteins (receptors)
CC that specifically bind to CCYPR, and in drug screening methods to
CC identify compounds that modulate the activity of CCYPR. CCYPR
CC nucleotides can be used to generate transgenic animal models of human
CC disease, and can be used in gene therapy in target cells with genetic
CC abnormalities with respect to the expression of CCYPR for the
CC treatment or prevention of a disorder associated with CCYPR.
CC Diseases which can be diagnosed, treated and prevented using CCYPR
CC proteins, nucleic acids, agonists or antagonists include immune,
CC developmental and cell signalling disorders, and cell proliferative
CC disorders including cancer. Specific examples of these disorders
CC include anaemia, epilepsy, arteriosclerosis, asthma, cancer, allergies,
CC diabetes mellitus, disorders of the menstrual cycle and infections
CC caused by bacteria.
XX
XX Sequence 580 BP; 167 A; 135 C; 160 G; 118 T; 0 other;
SQ
Query Match 2.9%; Score 23; DB 22; Length 580;
Best Local Similarity 100.0%; Pred. No. 0.74; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0;
Qy 645 TAAATGCCAGAGCAGGTGAAG 667
Db 520 TAAATGCCAGAGCAGGTGAAG 542
RESULT 24
AAD14983
ID AAD14983 standard; DNA: 611 BP.
AC AAD14983;
XX
XX 01-NOV-2001 (first entry)
DT
XX Human NOV4 INA.
DE
XX Human; NOV4: G-antigen; GAGE-like protein; interferon;
KW G-protein coupled receptor; GPCR; hepatocyte nuclear factor;
KW mast cell protease; gene therapy; proliferative disorder; cancer;
KW immune disorder; hepatic disorder; cirrhosis; viral infection;
KW hepatitis; neuroofactory system-related disorder; neurological disorder;
KW Parkinson's disease; infertility; autoimmune disease; arthritis;
KW multiple sclerosis; allergy; wound healing; cytostatic; nootropic;
KW immunosuppressive; neuroprotective; vulnery; hepatotropic; ds.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 1..173
FT 5'UTR /*tag= a
FT CDS 174..521
FT /*tag= b
FT /*product= "Human NOV4 protein"
FT 3'UTR 522..611
FT /*tag= c
XX
XX WO200161009-A2.
XX
XX 23-AUG-2001.
PD
XX 15-FEB-2001; 2001WO-US04828.
PF
XX 15-FEB-2000; 2000US-0182723.
PR 15-FEB-2000; 2000US-0182724.
PR 15-FEB-2000; 2000US-0182733.
PR 22-FEB-2000; 2000US-0183896.
PR 23-FEB-2000; 2000US-0184275.
PR 23-FEB-2000; 2000US-0184482.
PR 23-FEB-2000; 2000US-0184497.
PR 24-FEB-2000; 2000US-0184744.

PR 13-APR-2000; 2000US-0197083.
PR 10-AUG-2000; 2000US-0224157.
PR 18-SEP-2000; 2000US-0233405.
PR 27-SEP-2000; 2000US-0236060.
PR 02-JAN-2001; 2001US-0259414.
PR 18-JAN-2001; 2001US-0262454.
PR 14-FEB-2001; 2001US-0783429.
XX (CURA-) CURAGEN CORP.
PA Malyankar UM, Tchernev VT, Padigar M, Taupier RJ, Spytek KA;
PI Majumder K, Guo X, Spaderna SK, Boldog FL;
XX
XX WPI: 2001-514775/56.
DR P-PSDB: AAE08583.
DR
XX
XX Isolated novel polypeptides useful for diagnosis of and treating
PT cancer, infertility, autoimmune diseases, arthritis, multiple
PT sclerosis, allergies, wound healing and hepatic disorders -
XX
XX Claim 9; Page 14; 140pp; English.
PS
XX The present sequence is a human NOV4 DNA. The NOVX protein has homology
CC with one of G-antigen (GAGE)-like protein, interferon, G-protein coupled
CC receptor (GPCR), hepatocyte nuclear factor or mast cell protease. The
CC NOVX is useful for treating or preventing a pathology associated with
CC NOVX. It is also useful for determining the presence or amount of NOVX
CC DNA in a sample, for identifying a potential therapeutic agent and in
CC gene therapy. It is also useful for determining the presence of or
CC predisposition to a disease associated with altered levels of NOVX. It is
CC also useful for the diagnosis and treatment of proliferative disorders,
CC e.g., cancer, immune disorders, hepatic disorders, e.g., cirrhosis, viral
CC infections, e.g., hepatitis, neuroendocrine system-related disorders,
CC neurological disorders, e.g., Parkinson's disease, infertility,
CC autoimmune diseases, arthritis, multiple sclerosis, allergies and wound
CC healing.
XX
XX Sequence 611 BP; 196 A; 127 C; 152 G; 135 T; 1 other;
SQ
Query Match 2.9%; Score 23; DB 22; Length 611;
Best Local Similarity 100.0%; Pred. No. 0.74;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 645 TAAATGCCAGAGCAGGTGAAG 667
DB 472 TAAATGCCAGAGCAGGTGAAG 494
RESULT 25
AAK52902/c
ID AAK52902 standard; cDNA; 665 BP.
XX
XX AAK52902;
AC
XX
XX 06-NOV-2001 (first entry)
DT
XX
XX Human polynucleotide SEQ ID NO 2431.
DE
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200157190-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 05-FEB-2001; 2001WO-US04098.
PF
XX
XX 03-FEB-2000; 2000US-0496914.
PR
XX 27-APR-2000; 2000US-0560875.
PR
XX 20-JUN-2000; 2000US-0598075.
PR

PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 15-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
XX WPI: 2001-476283/51.
DR P-PSDB: AAM79769.
DR
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
PT
XX
XX Claim 1; Page 4711; 6221pp; English.
PS
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
XX Sequence 665 BP; 156 A; 170 C; 139 G; 200 T; 0 other;
SQ
Query Match 2.9%; Score 23; DB 22; Length 665;
Best Local Similarity 100.0%; Pred. No. 0.74;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 645 TAAATGCCAGAGCAGGTGAAG 667
DB 142 TAAATGCCAGAGCAGGTGAAG 120
RESULT 26
AAK51918
ID AAK51918 standard; cDNA; 673 BP.
XX
XX AAK51918;
AC
XX
XX 06-NOV-2001 (first entry)
DT
XX
XX Human polynucleotide SEQ ID NO 463.
DE
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200157190-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 05-FEB-2001; 2001WO-US04098.
PF
XX
XX 03-FEB-2000; 2000US-0496914.
PR
XX 27-APR-2000; 2000US-0560875.
PR
XX 20-JUN-2000; 2000US-0598075.
PR

PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728425.
XX (HYSE-) HYSEQ INC.
PA
XX
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
XX WPI: 2001-476283/51.
DR P-PSDB; AAM78785.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
XX Claim 1; Page 1680-1681; 622lpp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM60020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
XX Sequence 673 BP; 203 A; 145 C; 170 G; 155 T; 0 other;
Query Match 2.9%; Score 23; DB 22; Length 673;
Best Local Similarity 100.0%; Pred. No. 0.74;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 645 TAAATGCCAGACGAGGTGAAG 667
Db 536 TAAATGCCAGACGAGGTGAAG 558
XXXXXXXXXXXXXXXXXXXX
RESULT 27
ABL11924/c
ID ABL11924 standard; cDNA; 4652 BP.
XX
XX ABL11924;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ.ID NO 30254.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
KW
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
PN
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US09231.
PF
XX 23-MAR-2000; 2000US-191637P.
PR
XX 11-JUL-2000; 2000US-0614150.
PR
XX (PEKE) PE CORP NY.
PA
XX Venter JC, Adams M, Li PWD, Myers EW;
PI

XX WPI: 2001-656860/75.
DR P-PSDB; ABB67821.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Claim 1; SEQ ID NO 30254; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pt_sequences.
XX
XX Sequence 4652 BP; 1051 A; 1138 C; 1165 G; 1298 T; 0 other;
Query Match 2.7%; Score 21; DB 23; Length 4652;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 393 ACACACACAAACACAGAACCA 413
Db 3904 ACACACACAAACACAGAACCA 3884
XXXXXXXXXXXXXXXXXXXX
RESULT 28
AAQ66926/c
ID AAQ66926 standard; DNA; 83 BP.
XX
XX AAQ66926;
XX
XX 25-JAN-1995 (first entry)
XX
XX PolydA detection probe containing flourophores.
XX
XX Target sequence; binding ability; hairpin forming; probes;
KW imperfect hairpin; acceptor label moiety; donor moiety label;
KW fluorescence; fluorophores; specificity; base pair mismatches;
KW competitive arms; ss.
XX
XX Synthetic.
XX
XX Key modified_base 1 Location/Qualifiers
FT /*tag= a "Three-site aminoalkyl-derivatised
FT /note= oligonucleotide with a 6-carbon linker arm
FT attached to an additional fluorophore molecule"
FT stem_loop 1..41
FT /*tag= b
FT /note= "Forms imperfect hairpin with the competitive
FT arm containing 4 mismatches"
FT misc_difference 42
FT /*tag= c
FT /note= "6-carbon molecular linker containing a
FT fluorescein molecule and occupying the same
FT space as a base"
FT stem_loop 43..83
FT /*tag= d
FT /note= "Forms imperfect hairpin with the competitive
FT arm containing 4 mismatches"
FT modified_base 83
FT /*tag= e
FT /note= "Three-site aminoalkyl-derivatised
FT oligonucleotide with a 6-carbon linker arm

Thu Oct 17 09:19:16 2002

us-09-658-824-808.oligo.rng

attached to an additional fluorophore molecule"

FT XX
PN EP601889-A.
XX 15-JUN-1994.
PD XX
PF 10-DEC-1993; 93EP-0310007.
XX 10-DEC-1992; 92US-0990298.
XX (MAIN-) MAINE MEDICAL CENT RES INST.
PA Bagwell BC;
XX WPI; 1994-185245/23.
XX Nucleic acid probe for use in DNA or RNA hybridisation assays -
PT comprises a nucleotide sequence which is capable of forming one
PT or more hairpins
XX Disclosure; Fig 4; 25pp; English.
XX This sequence represents a hairpin forming probe of the invention. The
CC probes comprise a segment complementary to the target nucleotide and are
CC capable of forming at least 1 imperfect hairpin. The probes contain at
CC least one acceptor label moiety and at least one donor moiety label
CC which are covalently attached to the nucleotide sequence so that when
CC the hairpins are formed the moieties are in close proximity to allow
CC resonance energy transfer between them. This causes a reduction in the
CC fluorescence of the two fluorophores. The target sequence contains at
CC least 22 nucleotides to ensure specificity and avidity of the probe-
CC target hybridisation. The hairpins formed are imperfect hairpins and
CC the intentional base pair mismatches are introduced into the competitive
CC arms such that the specificity sequence will favour binding to the
CC target sequence. Therefore when the probe interacts with the target
CC sequence the competitive arm is displaced, increasing the distance
CC between the fluorophores, resulting in a change in fluorescent emission.
XX Sequence 83 BP; 26 A; 4 C; 9 G; 43 T; 1 other;
SQ
Query Match 2.6%; Score 20; DB 15; Length 83;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 762 CTTTCACCAAAAAAAAAA 781
Db 70 CTTTCACCAAAAAAAAAA 51
RESULT 29
ABAI7928
ID ABAI7928 standard; DNA; 21082 BP.
XX
AC ABAI7928;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 10259.
XX
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
KW antiparkinsonian; antisickling; antianemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX Homo sapiens.
OS
XX WO200159063-A2.
PN
XX 16-AUG-2001.
PD
XX

PF 17-JAN-2001; 2001WO-US01334.
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.

PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 18-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239035.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.

PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-541565/60.

Nucleic acids encoding 3224 human nervous system antigen polypeptides, useful for preventing, diagnosing and/or treating nervous system cancers and metastases.

Disclosure; SEQ ID NO 10258; 1701pp + Sequence Listing; English.

The invention relates to novel genes (AB11004-ABA21534) and proteins (AB114678-AB118001) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune

CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 21087 BP; 6131 A; 4253 C; 4393 G; 6310 T; 0 other;
Query Match 2.6%; Score 20; DB 22; Length 21087;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 762 CTTTCACCAAAAAA 781
|||||
Db 2635 CTTTCACCAAAAAA 2654
RESULT 31
AAS88056
ID AAS88056 standard; cDNA; 372 BP.
XX
AC AAS88056;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #23860.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639367/73.
DR P-PSDB; ABG23869.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 23860; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: the sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 372 BP; 126 A; 76 C; 82 G; 88 T; 0 other;
Query Match 2.4%; Score 19; DB 23; Length 372;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 763 TTTCACCAAAAAA 781
|||||
Db 347 TTTCACCAAAAAA 365
RESULT 32
AAK63150/C
ID AAK63150 standard; cDNA; 513 BP.
XX
AC AAK63150;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:8210.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ss.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.

PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0233397.
PR 14-SEP-2000; 2000US-0233398.
PR 14-SEP-2000; 2000US-0233399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.

PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249254.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-483426/52.

DR P-PSDB; AAM90369.

XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -

XX Claim 1; SEQ ID NO 8210; 3071pp + Sequence Listing; English.

XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.

SQ Sequence 513 BP; 155 A; 87 C; 138 G; 127 T; 6 other;

Query Match 2.4%; Score 19; DB 22; Length 513;

Best Local Similarity 100.0%; Pred. No. 55;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 763 TTTCACCAAAAAAAAAA 781

|||||

DB 20 TTTCACCAAAAAAAAAA 2

```

RESULT 33
AAZ80590/C
ID AAZ80590 standard; cDNA; 705 BP.
XX
XX AC AAZ80590;
XX DT 07-APR-2000 (first entry)
XX DE Human colon cancer cell line SW480 cDNA clone SEQ ID NO:674.
XX KW Human; gene expression product; diagnosis; tumour; colon cancer;
XX KW colorectal adenocarcinoma; cell line SW480; cell proliferation;
XX KW cytostatic; sarcoma; breast cancer; neoplasia; dysplasia;
XX KW hyperplasia; ds.
XX OS Homo sapiens.
XX PN W09964576-A2.
XX PD 16-DEC-1999.
XX PF 09-JUN-1999; 99WO-IB01062.
XX PR 10-JUN-1998; 98US-0088801.
XX PA (FARB ) BAYER CORP.
XX PI Endege WO, Steinmann KE, Astle JH, Burgess CC, Bushnell SE;
XX PI Carroll E, Catino TJ, Derti A, Ford DM, Lewis ME, Monahan JE;
XX PI Schlegel R;
XX DR WPI; 2000-087220/07.
XX PT Novel nucleic acids, used to develop products for the diagnosis and
XX PT treatment of disorders involving unwanted cell proliferation,
XX PT particularly cancers, especially colon cancer -
XX PS Claim 15; Page 399; 469pp; English.
XX CC AAZ79917 to AAZ80766 represent double stranded cDNA clones isolated from
XX CC the human colorectal adenocarcinoma (colon cancer) cell line SW480. The
XX CC cDNA clones can be used to generate antisense oligonucleotides which
XX CC can be used for antisense therapy. Methods and products from the present
XX CC invention can be used for identifying and/or classifying cancerous cells
XX CC present in a human tumour, particularly in solid tumours, e.g.
XX CC carcinomas and sarcomas, e.g. breast or colon cancers. The cDNA clones
XX CC can be used for developing agents for the diagnosis and treatment of
XX CC disorders involving unwanted cell proliferation, such as neoplasia,
XX CC dysplasia or hyperplasia.
XX SQ Sequence 705 BP; 144 A; 157 C; 166 G; 213 T; 25 other:

Query Match 2.4%; Score 19; DB 21; Length 705;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 763 TTTCACCAAAAAAAAAAAAA 781
Db 37 TTTCACCAAAAAAAAAAAAA 19

RESULT 34
AAQ32851/C
ID AAQ32851 standard; cDNA; 1547 BP.
XX
XX AC AAQ32851;
XX DT 05-MAY-1993 (first entry)
XX DE BMP2.
XX KW Bone; morphogenetic; protein; BMP; growth; vitamin D; systemic;
XX KW treatment; dimer; ss.

```

```

XX OS Rattus rattus.
XX PN W09221365-A.
XX PD 10-DEC-1992.
XX PF 26-MAY-1992; 92WO-US04356.
XX PR 05-JUN-1991; 91US-0709621.
XX PR 27-MAR-1992; 92US-0856110.
XX PA (PROC ) PROCTER & GAMBLE CO.
XX PI Stone RL;
XX DR WPI; 1992-433371/52.
XX PT Synergistic compsn. for generating mammalian bone growth -
XX PT comprises vitamin-D cpd. and bone morphogenetic protein
XX PS Disclosure; Page 27-29; 44pp; English.
XX CC The sequences given in AAQ32850-56 encode bone morphogenetic proteins
XX CC (BMP). BMP's increase bone growth and when used in conjunction with
XX CC vitamin D the level of new bone growth is greater than when a BMP or
XX CC vitamin D are used alone. The BMP's are administered for systemic
XX CC treatment at a dose range of 1pg to 100 microg. BMP are active as
XX CC dimers.
XX SQ Sequence 1547 BP; 377 A; 423 C; 410 G; 337 T; 0 other:

Query Match 2.4%; Score 19; DB 13; Length 1547;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 263 GGCCGAGGAGTGAGGGGC 281
Db 119 GGCCGAGGAGTGAGGGGC 101

RESULT 35
AAV01679/C
ID AAV01679 standard; cDNA; 1547 BP.
XX
XX AC AAV01679;
XX DT 27-MAR-1998 (first entry)
XX DE Bone morphogenetic protein BMP-2 encoding DNA.
XX KW Bone morphogenetic protein; BMP; growth; vitamin D; fracture;
XX KW arthritis; surgical lesion; periodontal disease; osteoporosis;
XX KW rickets; ds.
XX OS Unidentified.
XX PN US5693615-A.
XX PD 02-DEC-1997.
XX PF 05-JUN-1991; 91US-0709621.
XX PR 07-SEP-1993; 93US-0117367.
XX PR 05-JUN-1991; 91US-0709621.
XX PR 27-MAR-1992; 92US-0856110.
XX PR 09-DEC-1992; 92US-0988363.
XX PR 13-MAY-1994; 94US-0243435.
XX PR 23-JAN-1995; 95US-0377292.
XX PA (PROC ) PROCTER & GAMBLE CO.
XX PI Stone RL;

```

XX WPI: 1998-031788/03.
XX Generation of new bone growth - by co-administering bone
PT morphogenetic protein and vitamin D
XX
XX Claim 1; Column 19-22; 18pp; English.
XX
CC A new method has been developed for generating new bone growth in a
CC mammal. The method comprises administering a bone morphogenetic protein
CC in combination with a vitamin D compound, where: (a) the bone
CC morphogenetic protein is BMP-2 and is administered in an amount of
CC 500-1000 ng in combination with about 6 ng vitamin D compound; or (b)
CC the bone morphogenetic protein is BMP-4 and is administered in an
CC amount of about 62.5 ng in combination with about 6 ng vitamin D
CC compound. The present sequence encodes BMP-2. The method is used for
CC treating bone defects or disorders, e.g. fractures, surgical lesions,
CC periodontal disease, osteoporosis, arthritis and rickets.
XX
SQ Sequence 1547 BP; 377 A; 423 C; 410 G; 337 T; 0 other;

Query Match 2.4%; Score 19; DB 19; Length 1547;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 263 GGCCGAGGAGTGGAGGGC 281
Db 119 GGCCGAGGAGTGGAGGGC 101

RESULT 36
AAN80632/c
ID AAN80632 standard; DNA; 1606 BP.
XX
AC AAN80632;
XX
DT 08-OCT-1990 (first entry)
XX
DE Human Bone Morphogenic Protein-2 class I cDNA.
XX
KW Bone morphogenic protein; hBMP-2 class I; probes; cartilage formation;
KW bone formation; osteogenic cpds; prodontal disease; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 356..1543
FT /*tag= a
FT /product=hBMP-2 class I
XX
PN W08800205-A.
XX
XX 14-JAN-1988.
XX
XX 30-JUN-1987; 87WO-0501537.
XX
XX 26-MAR-1987; 87US-0031346.
XX
XX (GENE-) GENETICS INST INC.
XX
XX Wozney JW, Rosen VA;
XX
XX WPI: 1988-021565/03.
XX P-PSDB; AAP80619.
XX
XX Bone morphogenic proteins - obtd. using recombinant DNA and used
PT for inducing cartilage and bone formation.
XX
XX Disclosure; 7pp; English.
XX
CC The HindIII-SacI bovine genomic bBMP-2 fragment described in AAN80627
CC is subcloned into M13, labelled and used as probe to screen
CC polyadenylated RNAs from various cells and tissue sources.
CC

CC Sequence analysis of the strongly hybridising clones hBMP-2 class I
CC (-BMP-2) indicated that they have extensive homology with the sequence
CC given in AAN80622.
CC The partial sequence is compiled from lambda U20S-39 and several
CC other hBMP-2 class I cDNA recombinants.
CC This human cDNA hBMP-2 class I contains an open reading frame of 1188
CC bp, encoding a protein of 396 amino acids. The protein is preceded by a
CC 5' untranslated region of 342 bp with stop codons in all frames. The 13
CC bp region preceding this 5' untranslated region represents a linker used
CC in the cDNA cloning procedures.
CC See also AAN80619-N80636 and AAN81963-64.
XX
SQ Sequence 1606 BP; 399 A; 430 C; 423 G; 354 T; 0 other;

Query Match 2.4%; Score 19; DB 9; Length 1606;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 263 GGCCGAGGAGTGGAGGGC 281
Db 151 GGCCGAGGAGTGGAGGGC 133

RESULT 37
AAQ14036/c
ID AAQ14036 standard; DNA; 1607 BP.
XX
AC AAQ14036;
XX
DT 13-JAN-1992 (first entry)
XX
DE Human BMP-2A in lambda U20S-39.
XX
KW Bone; cartilage; osteoinductive protein; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 356..1546
FT /*tag= a
FT Misc_RNA 356..1543
FT /*tag= b
FT /label= claim 1(b)
XX
PN US5013649-A.
XX
XX 07-MAY-1991.
XX
XX 08-APR-1988; 88US-0179100.
XX
XX 01-JUL-1986; 86US-0880776.
XX 04-APR-1988; 86US-0179100.
XX
XX (GENE-) GENETICS INST INC.
XX
XX WPI: 1991-309401/42.
XX P-PSDB; AAR14241.
XX
XX New DNA sequences encoding osteo-inductive protein - useful for
PT stimulating bone and cartilage re formation e.g. for wound healing
PT and tissue repair.
XX
XX Disclosure; Fig 2; 20pp; English.
XX
XX The sequence was obtd. from clone lambda U20S-39 (ATCC 40345)
CC which was isolated from a cDNA library prepd. using human cell line
CC U-2 OS RNA. It is one of two classes of clones isolated
CC distinguished by hybridisation characteristics. This class,
CC designated hBMP-2A (previously BMP-2 and BMP-2 Class I) hybridised
CC strongly with the probe which was derived from the bovine BMP-2A
CC sequence, and showed strong homology with that sequence. The second
CC class, hBMP-2B (previously BMP-4 and BMP-2 Class II) hybridised only
CC weakly and is less homologous, esp. at the 3' end. The gene prod.

CC is an osteoinductive protein useful for inducing bone/cartilage
CC repair, wound healing and tissue repair. Typical applications
CC include healing of bone fractures; improved fixation of artificial
CC joints; in cosmetic plastic surgery; and in treatment of periodontal
CC disease, burns, incisions, ulcers etc.
CC See also AAQ14035 and AAQ14037, AAQ14909 and AAQ14910.
XX
SQ Sequence 1607 BP; 400 A; 430 C; 421 G; 356 T; 0 other;
Query Match 2.4%; Score 19; DB 12; Length 1607;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 263 GCCCGAGGAGTGGAGGGC 281
Db 151 GCCCGAGGAGTGGAGGGC 133
RESULT 38
AAQ31869/c
ID AAQ31869 standard; DNA: 1607 BP.
XX
AC AAQ31869;
XX
DT 16-APR-1993 (first entry)
XX
DE Human BMP-2A sequence.
XX
KW Bone Morphogenic Protein; bacteriophage lambda U2OS-39; ATCC 40345;
KW bone fracture; cartilage defect; osteoporosis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 356..1546
FT FT /*tag= a
FT mat_peptide 1202..1543
FT FT /*tag= b
FT FT /note= "preceded by pre-pro region"
XX
PF US5166058-A.
XX
PD 24-NOV-1992.
XX
PF 11-JUL-1989; 89US-0378537.
XX
PR 01-JUL-1986; 86US-0880776.
PR 17-DEC-1986; 86US-0943332.
PR 20-MAR-1987; 87US-0028285.
PR 08-APR-1988; 88US-0179100.
PR 11-JUL-1989; 89US-0378537.
XX
PA (GEMY) GENETICS INST INC.
XX
PI Rosen VA, Wang EA, Wozney JM;
XX
XX WPI; 1992-414955/50.
DR P-PSDB; AAR29281.
XX
PT DNA encoding osteo-inductive proteins - used for producing BMP-2A
PT and BMP-2B for inducing bone or cartilage formation and wound
PT healing
XX
PS Claim 1; Fig 2; 22pp; English.
XX
CC A human U2OS cDNA library was screened with a bovine BMP-2A probe
CC (see AAQ31868). Sequence analysis of strongly hybridising clones
CC designated hBMP-2A (previously designated BMP-2 and BMP-2 Class 1)
CC indicated that they have extensive homology with the partial bovine
CC BMP-2A sequence; weakly hybridising clones were designated hBMP-2B
CC (previously designated BMP-4 and BMP-2 Class II) and sequence
CC analysis indicated these clones were quite homologous with the

CC bovine sequence at their 3' end but less so at the 5' end. A
CC full-length human BMP-2A cDNA clone was obtained by screening the
CC U2OS library with the insert of a BMP-2B subclone.
XX
SQ Sequence 1607 BP; 400 A; 430 C; 422 G; 355 T; 0 other;
Query Match 2.4%; Score 19; DB 13; Length 1607;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 263 GCCCGAGGAGTGGAGGGC 281
Db 151 GCCCGAGGAGTGGAGGGC 133
RESULT 39
AAQ41291/c
ID AAQ41291 standard; DNA: 1607 BP.
XX
AC AAQ41291;
XX
DT 13-SEP-1993 (first entry)
XX
DE Human BMP-2 gene.
XX
KW Bone morphogenetic protein; bone defect treatment; healing; wound;
KW injury; tissue repair; osteoporosis; burns; incisions; ulcers;
KW neuronal survival increase; fracture reduction; cartilage growth;
KW induction; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 356..1543
FT FT /*tag= a
FT WO9309229-A.
XX
PD 13-MAY-1993.
XX
PF 02-NOV-1992; 92WO-US09430.
XX
PR 04-NOV-1991; 91US-0787496.
PR 07-APR-1992; 92US-0864692.
XX
PA (GEMY) GENETICS INST INC.
XX
PI Israel D, Wolfman NM;
XX
XX WPI; 1993-167696/20.
DR P-PSDB; AAR36732.
XX
PT Recombinant hetero-dimeric BMP proteins - are useful in treating
PT bone defects, healing bone injury and in wound healing
XX
PS Disclosure; Fig 1; 169pp; English.
XX
CC The sequence is that encoding the human bone morphogenetic protein
CC BMP-2. It may be used in the prodn. of a recombinant heterodimeric
CC protein having bone stimulating activity. This heterodimer is
CC encoded by a sequence encoding BMP-2 or a fragment and a sequence
CC encoding a second protein or fragment selected from BMP-5, BMP-6,
CC BMP-7 or BMP-8. It may be used in compsns. for wound healing, tissue
CC repair, and in similar compsns. which have been indicated for the
CC use of individual BMPs. Increased potency of the heterodimer over
CC individual BMPs may permit lower dosages to be administered. A
CC heterodimeric protein which induces cartilage and/or bone growth
CC in circumstances where bone is not normally formed, has
CC applications in the healing of bone fractures and cartilage defects
CC in humans and other animals. The heterodimer may have prophylactic use
CC in closed as well as open fracture reduction and also in the
CC improved fixation of artificial joints. De novo bone formation
CC induced by an osteogenic agent contributes to the repair of

CC congenital, trauma induced or oncologic resection induced craniofacial
 CC defects, and also is useful in cosmetic plastic surgery. It may be
 CC used in the treatment of periodontal disease and in other tooth
 CC repair processes. It may also be useful in the treatment of
 CC osteoporosis, wound healing (e.g. burns, incisions and ulcers) and
 CC related tissue repair, and may increase neuronal survival and be useful
 CC in the transplantation and treatment of conditions exhibiting a decrease
 CC in neuronal survival. It may be combined with other agents beneficial to
 CC the bone and/or cartilage defect, wound or tissue in question, e.g. EGF,
 CC PDGF, TGF-alpha, TGF-beta and insulin-like growth factor
 XX
 SQ Sequence 1607 BP; 400 A; 430 C; 423 G; 354 T; 0 other;
 Query Match 2.4%; Score 19; DB 14; Length 1607;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 263 GCCCGAGGAGTGGAGGGGC 281
 |||||
 Db 151 GCCCGAGGAGTGGAGGGGC 133
 RESULT 40
 AAT78941/c
 ID AAT78941 standard; cDNA; 1607 BP.
 XX
 AC AAT78941;
 XX
 DT 22-JAN-1998 (first entry)
 XX
 DE Human bone morphogenic protein (BMP) 2A cDNA.
 XX
 KW bone morphogenic protein; BMP 2A; cartilage; periodontal disease;
 KW tissue repair; osteoporosis; treatment; oligonucleotide probe; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 5'UTR 1..355
 FT /*tag= a
 FT CDS 356..1546
 FT /*tag= b
 FT misc_feature 356..424
 FT /*tag= c
 FT /*note= "pre-portion of BMP-2A"
 FT misc_feature 425..1201
 FT /*tag= d
 FT /*note= "pro-portion of BMP-2A"
 FT mat_peptide 1202..1543
 FT /*tag= e
 FT 3'UTR 1547..1607
 FT /*tag= f
 XX
 PN US5631142-A.
 XX
 PD 20-MAY-1997.
 XX
 PF 07-SEP-1993; 93US-0118363.
 XX
 PR 11-JUL-1989; 89US-0378537.
 PR 17-DEC-1986; 86US-0943332.
 PR 20-MAR-1987; 87US-0028285.
 PR 08-APR-1988; 88US-0179100.
 PR 18-MAY-1992; 92US-0884353.
 PR 07-SEP-1993; 93US-0118363.
 PR 01-JUL-1986; 86US-0860776.
 XX
 PA (GEM) GENETICS INST INC.
 XX
 PI Rosen VA, Wang EA, Wozney JM;
 XX
 DR WPI; 1997-288573/26.

DR P-PSDB; AAW24849.
 XX Production of human bone morphogenic protein 2A or 2B in cell
 PT culture - useful inducing bone or cartilage production, in wound
 PT healing and tissue repair
 XX
 PS Claim 2; Fig 2; 22pp; English.
 XX This cDNA sequence encodes the human bone morphogenic protein (BMP) 2A.
 CC Oligonucleotide probes were synthesised based on a partially sequenced
 CC BMP-2A protein isolated from ground bovine powder. The probes were used
 CC to screen a bovine liver DNA library to obtain the BMP-2A encoding DNA
 CC sequence. The DNA was used to screen a U-2 OS human cell line cDNA
 CC library to obtain this human BMP-2A cDNA. BMPs can be used to induce
 CC bone and cartilage formation, and in wound healing and tissue repair.
 CC They can also be used for treating periodontal disease or osteoporosis.
 XX
 SQ Sequence 1607 BP; 400 A; 430 C; 422 G; 355 T; 0 other;
 Query Match 2.4%; Score 19; DB 18; Length 1607;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 263 GCCCGAGGAGTGGAGGGGC 281
 |||||
 Db 151 GCCCGAGGAGTGGAGGGGC 133
 Search completed: October 16, 2002, 22:03:13
 Job time : 243 secs

10	305	39.1	391	22	AAS38132	Novel human diago
11	226	28.9	300	20	AA55597	Human cDNA clone F
12	216.4	27.7	659	22	AA158744	Human polynucleoti
13	214.8	27.5	673	22	AAK51918	Human polynucleoti
14	201.8	25.8	611	22	AAD14983	Human NOV4 DNA. H
15	201.2	25.8	503	23	AA569484	DNA encoding novel
16	197	25.2	208	21	AA17098	Human secreted pro
17	195.6	25.0	532	22	AA160530	Human polynucleoti
18	174.8	22.4	201	21	AAA42613	DNA encoding novel
19	171.8	22.0	661	23	AA569486	Human cell cycle a
20	169	21.6	580	22	AAF59637	Human metastatic m
21	163.8	21.0	1092	21	AAA45936	Human polynucleoti
22	142.6	18.3	665	22	AAK52502	DNA encoding novel
23	139.2	17.8	750	23	AA591235	DNA encoding novel
24	135	17.3	320	22	AAH83260	Human ovarian tumo
25	132.2	17.1	214	22	AA524637	Human ovarian PCR-
26	132.6	17.0	2182	21	AA5259112	Human secreted pro
27	111.6	14.3	539	19	AAV18721	cDNA encoding GAGE
28	110	14.1	540	20	AAK90521	GAGE-4 tumour reje
29	110	14.1	540	20	AAK90523	GAGE-6 tumour reje
30	108.4	13.9	532	19	AAV18720	cDNA encoding GAGE
31	108.4	13.9	532	20	AAK90522	GAGE-5 tumour reje
32	108.4	13.9	535	19	AAV18717	cDNA encoding GAGE
33	108.4	13.9	538	20	AAK90519	GAGE-2 tumour reje
34	108.4	13.9	560	19	AAV18718	cDNA encoding GAGE
35	108.4	13.9	560	20	AAK90520	GAGE-3 tumour reje
36	108.4	13.9	1245	22	AA560826	Human cancer agent
37	106.8	13.7	1024	21	AA297217	Human prostate can
38	106.4	13.6	475	22	AAD14981	Human NOV2 DNA. H
39	105.8	13.5	1024	21	AA297216	Human prostate can
40	103.2	13.2	1051	22	AAD14982	Human NOV3 DNA. H
41	102.4	13.1	190	21	AAK10552	Human secreted pro
42	99	12.7	541	19	AAV18719	cDNA encoding GAGE
43	95.8	12.3	530	22	AA560104	Human cancer agent
44	91.4	11.7	571	21	AA533316	Human secreted pro
45	89.6	11.5	472	21	AAZ95012	CSG Prol18 EST use

ALIGNMENTS

RESULT 1

AAF68861

ID AAF68861 standard; cDNA; 781 BP.

XX AAF68861;

AC

CC

DT 12-APR-2001 (first entry)

Human lung tumour protein related nucleotide sequence SEQ ID NO:808.

Human; lung cancer; lung tumour; lung tumour protein; gene therapy;

lung cancer antigen; lung tumour-specific antigen; diagnosis; vaccine;

cytostatic; antisense inhibition; ss.

Homo sapiens.

OS

XX

PN WO200100828-A2.

XX

PD 04-JAN-2001.

XX

PF 30-JUN-2000; 2000WO-US18061.

XX

PR 30-JUN-1999; 99US-0346492.

XX

PR 15-OCT-1999; 99US-0419356.

XX

PR 17-DEC-1999; 99US-0468667.

XX

PR 30-DEC-1999; 99US-0476300.

XX

PR 06-MAR-2000; 2000US-0519642.

XX

PR 22-MAR-2000; 2000US-0533077.

XX

PR 10-APR-2000; 2000US-0546259.

XX

PR 27-APR-2000; 2000US-0560406.

XX

PR 05-JUN-2000; 2000US-0589184.

XX

PA (CORI-) CORIXA CORP.
 XX Wang T, Banquer CS, Lodes MJ, Fanger GR, Vedwick TS, Carter D;
 PI Retter MW, Mannion J;
 XX WPI: 2001-071498/08.
 XX Lung tumor-associated proteins and the nucleic acids that encode them,
 PT useful for preventing, diagnosing and treating lung cancer .
 XX Claim 4; Page 426; 436pp; English.
 XX The present invention describes immunogenic portions of lung tumour-
 CC associated proteins (I) and the nucleic acids (NAS) that encode them.
 CC (I) have cytostatic activity and can be used in gene therapy, antisense
 CC inhibition and in vaccines. The NAS and the lung tumour-associated
 CC proteins they encode may be used in the prevention, treatment and
 CC diagnosis of diseases associated with their inappropriate expression,
 CC especially lung cancers. For example, the NAS may be administered to
 CC treat diseases by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of the protein by expressing inactive proteins
 CC or to supplement the patient's own production of (I). Additionally, the
 CC NAS may be used to produce the lung-tumour associated protein, according
 CC to standard recombinant DNA methodology. Conversely, antisense NA
 CC molecules may be administered to down regulate protein expression by
 CC binding with the cells own genes and preventing their expression. The NA
 CC and complementary sequences may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar NA sequences in
 CC samples, and hence which patients may be in need of treatment for lung
 CC cancer. The (I) may be used as antigens in the production of antibodies
 CC and in assays to identify modulators (agonists and antagonists) of the
 CC expression and activity of the protein. AAF68083 to AAF68878 and
 CC AAB76848 to AAB76878 represent human lung tumour protein related
 CC nucleotide and protein sequences which are used in the exemplification
 CC of the present invention.
 XX Sequence 781 BP: 230 A; 181 C; 234 G; 136 T; 0 other;
 SQ

Query Match 100.0%; Score 781; DB 22; Length 781;
 Best Local Similarity 100.0%; Pred. No. 6e-224;
 Matches 781; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCGGAGCTGTGAGCGGCGGACTCGGCTCCCTGAGGTCTGGATCTTTCTCCGCTACT 60
 Db 1 GCGGCGGAGCTGTGAGCGGCGGACTCGGCTCCCTGAGGTCTGGATCTTTCTCCGCTACT 60

Qy 61 GAGACACIAGCGGGTAGGTTCACAGGACAGATCCAACTGGGAGTTGAAGTGTGAGTGAGCT 120
 Db 61 GAGACACGCGGGTAGGTTCACAGGACAGATCCAACTGGGAGTTGAAGTGTGAGTGAGCT 120

Qy 121 GAAGAAGAACAGGAGGTTCCGAGGAGGTTGTGTGTCAGTCACTCAGAGTGAGAAGGCC 180
 Db 121 GAAGAGGAACAGGAGGTTCCGAGGAGGTTGTGTGTCAGTCACTCAGAGTGAGAAGGCC 180

Qy 181 CTCGAAGTCGTCCTCTCTCATCGGTGCGACGCCCATGAGCTTCTTCTCTCTCTGTCACG 240
 Db 181 CTCGAAGTCGTCCTCTCTCATCGGTGCGACGCCCATGAGCTTCTTCTCTCTCTGTCACG 240

Qy 241 GCCATAACTAGGAGGAGGAGGCGCGAGGAGTGGAGGGCTCAGCGCAAGCTGGGGTGC 300
 Db 241 GCCATAACTAGGAGGAGGAGGCGCGAGGAGTGGAGGGCTCAGCGCAAGCTGGGGTGC 300

Qy 301 TGTGGGGGTATCCGAGTCCAGAGCACTTGGAAACCCCGACAGAGATTTCTGACTCCC 360
 Db 301 TGTGGGGGTATCCGAGTCCAGAGCACTTGGAAACCCCGACAGAGATTTCTGACTCCC 360

Qy 361 CAGAGCGGACAGGAGGAGGCGGATGAGCGACACACACAAACACAGACACACAGCC 420
 Db 361 CAGAGCGGACAGGAGGAGGCGGATGAGCGACACACACAAACACAGACACACAGCC 420

Qy 421 ACTCCAGAGCCCAAGTAATGGAGAGGCCCAAAAAGAGAACACAGAGCTGAAGTCGGG 480
 Db 421 AGTCCAGGAGCCCAAGTAATGGAGAGGCCCAAAAAGAGAACACAGAGCTGAAGTCGGG 480

Qy 481 ATCTTACACCTGGGCGAGCAGACAGAGAAGATCAGATACAGCTGAGATCCAGTCCGCG 540
 Db 481 ATCTTACACCTGGGCGAGCAGACAGAGAAGATCAGATACAGCTGAGATCCAGTCCGCG 540

Qy 541 ACATGGAAGGTGATCTGCAAGAGCTGCATCAGCTCAACACCGGGGATAAATCTGGATTG 600
 Db 541 ACATGGAAGGTGATCTGCAAGAGCTGCATCAGCTCAACACCGGGGATAAATCTGGATTG 600

Qy 601 GGTTCGGCGCTCAAGGTGAAGATAATACCTAAAGAGGAACTGTAAATGCGCAAGCA 660
 Db 601 GGTTCGGCGCTCAAGGTGAAGATAATACCTAAAGAGGAACTGTAAATGCGCAAGCA 660

Qy 661 GGTGAAGACCAACCAAGTTTAAATGAAGCAAGCTGAAACAGCAAGCTGGTTTAT 720
 Db 661 GGTGAAGACCAACCAAGTTTAAATGAAGCAAGCTGAAACAGCAAGCTGGTTTAT 720

Qy 721 ATTAGATATTTGACATTAACACTCTCAATAAAGTTTTCAGCTTTTACCAAAAAA 780
 Db 721 ATTAGATATTTGACATTAACACTCTCAATAAAGTTTTCAGCTTTTACCAAAAAA 780

Qy 781 A 781
 Db 781 A 781

RESULT 2
 AAS37109 standard; cDNA: 396 BP.
 AC AAS37109,
 DT 17-DEC-2001 (first entry)
 DE ~~NOVEL~~ human diagnostic and therapeutic gene #167.
 XX Human; cancer; breast; lung; colon; prostate; cytostatic; diagnostic; ss.
 XX Homo sapiens.
 PN WO200166753-A2.
 PD 13-SEP-2001.
 XX 09-MAR-2001; 2001WO-US07782.
 XX 09-MAR-2000; 2000US-0188609.
 XX (CHIR) CHIRON CORP.
 XX (HYSE-) HYSEQ INC.
 XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
 PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Kassam A, Lamson G;
 PI Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
 PI Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;
 XX WPI: 2001-530177/58.
 XX New polynucleotides and polypeptides, useful for diagnosis and
 treatment of breast, lung and colon cancer -
 Claim 1; Page 638; 1193pp; English.
 CC The invention relates to new polynucleotides and polypeptides, useful for
 diagnosis and treatment of breast, lung and colon cancer. The sequences
 can be used in detecting differentially expressed genes correlated with a
 cancerous state of a mammalian cell, comprising detecting at least one
 differentially expressed gene product in a test sample derived from a
 cell suspected of being cancerous. They can also be used to inhibit
 tumour growth by modulating expression of a gene product. AAS36943-
 AAS39338 represent novel human diagnostic and therapeutic coding
 sequences of the invention.
 CC

SQ Sequence 396 BP; 82 A; 97 C; 146 G; 71 T; 0 other;

Query Match 50.1%; Score 391; DB 22; Length 396;
Best Local Similarity 100.0%; Pred. No. 4.2e-107;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CGGCGAGCTGTAGCCGGGACTCGGGTCCCTGAGGCTCTGATTCCTTCTCCGCTACTG 61
DB 1 CGGCGAGCTGTAGCCGGGACTCGGGTCCCTGAGGCTCTGATTCCTTCTCCGCTACTG 60

QY 62 AGACACGGCGGTAGGTCCACAGGCAGATCCAACTGGGAGTTGAAGTGTGAGTGAGAGTG 121
DB 61 AGACACGGCGGTAGGTCCACAGGCAGATCCAACTGGGAGTTGAAGTGTGAGTGAGAGTG 120

QY 122 AGAGGAACACAGGCTCCCGAGGGTTGTGTGTGCTAGTCACTCAGAGTGAGAGAGGCC 181
DB 121 AAGAGGAACACAGGCTCCCGAGGGTTGTGTGTGCTAGTCACTCAGAGTGAGAGAGGCC 180

QY 182 TCGAAGTCGTCTCTCTCATGCGGTGCCAGGCCCATGGACCTTCTTGTCTCGTCACGG 241
DB 181 TCGAAGTCGTCTCTCTCATGCGGTGCCAGGCCCATGGACCTTCTTGTCTCTCGTCACGG 240

QY 242 CCATAACTAGGAGGAAGAGGGCCGAGGAGTGGAGGGGCTCAGCGGAAGCTGGGGTGCT 301
DB 241 CCATAACTAGGAGGAAGAGGGCCGAGGAGTGGAGGGGCTCAGCGGAAGCTGGGGTGCT 300

QY 302 GTTGGGGTATCCGAGTCCACAGACACTGGAACCCGACAGAGATTTCTGGACTCCCC 361
DB 301 GTTGGGGTATCCGAGTCCACAGAGCACCCTGGAACCCGACAGAGATTTCTGGACTCCCC 360

QY 362 AGACGGGACAGGAGAGGGGACGCGCATGAGCG 392
DB 361 AGACGGGACAGGAGAGGGGACGCGCATGAGCG 391

RESULT 3
AAH64751
ID AAH64751 standard; cdna; 762 BP.
XX
AC AAH64751;
XX
DT 11-SEP-2001 (first entry)
XX
DE Human secreted protein cDNA, SEQ ID NO: 27.
XX
KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KW GENSET; ss.
XX
OS Homo sapiens.
XX
PN WO200142451-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-1B01938.
XX
PR 08-DEC-1999; 99US-0169629.
PR 06-MAR-2000; 2000US-0187470.
XX
PA (GSET) GENSET.
XX
PI Dumas Milne Edwards J, Bougueleret L, Jobert S;
XX
DR WPI; 2001-367870/38.
DR P-PSDB; AAG89148.
XX
PT Full length GENSET human nucleic acids encoding potentially secreted
PT proteins, useful in gene therapy and vaccination against a variety of
PT diseases, and for diagnosis of those diseases -
XX
PS Claim 7; Page 586; 921pp; English.
XX
CC The invention relates to full length GENSET human nucleic acids encoding

CC potentially secreted proteins. The nucleic acids and the polypeptides
CC they encode may be used in the prevention, treatment and diagnosis of
CC diseases associated with inappropriate GENSET gene expression. For
CC example, they be used to treat disorders associated with decreased
CC GENSET gene expression by rectifying mutations or deletions in a
CC patient's genome that affect the activity of GENSET or by supplementing
CC the patient's own production of GENSET polypeptides. Conversely,
CC antisense nucleic acid molecules may be administered to down regulate
CC GENSET expression by binding with the cells' own genes and preventing
CC their expression. The sense and antisense nucleic acids may also be
CC used as DNA probes in diagnostic assays to detect and quantitate the
CC presence of similar nucleic acid sequences in samples, and hence to
CC determine which patients may be in need of restorative therapy.
CC The GENSET polypeptides may be used as antigens in the production of
CC antibodies and in assays to identify modulators (agonists and
CC antagonists) of GENSET polypeptide expression and activity. The
CC present sequence is a GENSET nucleic acid of the invention.
XX
SQ Sequence 762 BP; 220 A; 192 C; 186 G; 164 T; 0 other;

Query Match 50.1%; Score 391; DB 22; Length 762;
Best Local Similarity 98.7%; Pred. No. 5.9e-107;
Matches 394; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 383 GGCATGAGCGACACACACAAACACAGAGACCCAGTCCAGGAGCCCGCAATGG 442
DB 361 GACAGCGCGGACACACACAAACACAGAGACCCAGTCCAGGAGCCCGCAATGG 420

QY 443 AGAGCCCCAAAAGAGAGAACCCAGCAGCTGAAAGTCGGGATCTACACCTGGCGACGAC 502
DB 421 AGAGCCCCAAAAGAGAGAACCCAGCAGCTGAAAGTCGGGATCTACACCTGGCGACGAC 480

QY 503 AGAAGAAGTACAGGATACAGCTGAGATCCAGTCCAGTCCGCGACATGGAAGGTGATCTCAAGA 562
DB 481 AGAA''AGATCAG:ATACAGCTGAGATCCAGTCCGCGACATGGAAGGTGATCTCAAGA 540

QY 563 GCTGATCAGTCAACACCCGGGGATAAATCTGGATTTCGGGTTCCGGGCTCAAGGTGAAGA 622
DB 541 GCTGATCAGTCAACACCCGGGGATAAATCTGGATTTCGGGTTCCGGGCTCAAGGTGAAGA 600

QY 623 TAATACCTAAAGAGGAGCACTGTAAATGCCAGAGAGGTGAAGCAACACCAAGTTT 682
DB 601 TAATACCTAAAGAGGAGCACTGTAAATGCCAGAGAGGTGAAGCAACACCAAGTTT 660

QY 683 AAATGAAGACAGCTGAAACACGCAAGCTGGTTTTATATTAGATATTGACTTAACTA 742
DB 661 AAATGAAGACAGCTGAAACACGCAAGCTGGTTTTATATTAGATATTGACTTAACTA 720

QY 743 TCTCAATAAAGTTTTCAGCTTTCACCAAAAAA 781
DB 721 TCTCAATAAAGTTTTCAGCTTTCACCAAAAAA 759

RESULT 4
AAF68151
ID AAF68151 standard; cdna; 399 BP.
XX
AC AAF68151;
XX
DT 12-APR-2001 (first entry)
XX
DE Human lung tumour protein related nucleotide sequence SEQ ID NO:69.
XX
KW Human; lung cancer; lung tumour; lung tumour protein; gene therapy;
KW lung cancer antigen; lung tumour-specific antigen; diagnosis; vaccine;
KW cytostatic; antisense inhibition; ss.
XX
OS Homo sapiens.
XX
PN WO200100828-A2.
XX
PD 04-JAN-2001.
XX

CC expression and activity of the protein. AAF68083 to AAF68878 and
CC AAB76848 to AAB76878 represent human lung tumour protein related
CC nucleotide and protein sequences which are used in the exemplification
CC of the present invention.

XX New peptide useful as a marker for the diagnosis of breast cancer -
 XX
 XX
 XX Claim 1; Page 1118; 3695pp; English.
 XX
 CC The invention relates to human breast cancer expressed polynucleotides
 CC (AAU07544-AAU26789), and methods of assessing whether a patient is
 CC afflicted with breast cancer by examining the correlation between the
 CC expression of certain markers and the cancerous state of breast cells.
 CC The polynucleotides and encoded polypeptides are potential markers for
 CC detecting, diagnosing, monitoring, characterising treating and
 CC potentially preventing breast cancer. The polynucleotides and encoded
 CC polypeptides are also useful for isolating compounds with cytostatic
 CC activity.

QY 547 AAGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTGGTTCC 606
DB 244 AAGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTGGTTCC 303
QY 607 GCGTCAAGGTCAAGATAATACCTAAAGAGGAACACTGTAAATGCCAGAGCAGGTGAA 666
DB 304 GCGTCAAGGTCAAGATAATACCTAAAGAGGAACACTGTAAATGCCAGAGCAGGTGAA 363
QY 667 GAGCAACCAAGTTTAAATGAAGACAGCTGAAACAGCGCAAGCTGTTTATATTTA-G 725
DB 364 GAGCAACCAAGTTTAAATGAAGACAGCTGAAACAGCGCAAGCTGTTTATATTTAG 423
QY 726 ATATTGACTTAAACTATCTCAATAAAGTTTTCAGCTTTCACCAAAAAA 781
DB 424 ATATTGACTTAAACTATCTCAATAAAGTTTTCAGCTTTCACCAAAAAA 479
RESULT 9
AAF68852
ID AAF68852 standard; cDNA; 461 BP.
XX
AC AAF68852;
XX
DT 12-APR-2001 (first entry)
XX
DE Human lung tumour protein related nucleotide sequence SEQ ID NO:792.
XX
XX Human; lung cancer; lung tumour; lung tumour protein; gene therapy;
KW lung cancer antigen; lung tumour-specific antigen; diagnosis; vaccine;
KW cytostatic; antisense inhibition; ss.
XX
OS Homo sapiens.
XX
PN WO200100828-A2.
XX
PD 04-JAN-2001.
XX
XX 30-JUN-2000; 2000WO-US18061.
XX
XX 30-JUN-1999; 99US-0346492.
PR 15-OCT-1999; 99US-0419356.
PR 17-DEC-1999; 99US-0466867.
PR 30-DEC-1999; 99US-0476300.
PR 06-MAR-2000; 2000US-0519642.
PR 22-MAR-2000; 2000US-0533077.
PR 10-APR-2000; 2000US-0546259.
PR 27-APR-2000; 2000US-0560406.
PR 05-JUN-2000; 2000US-0589184.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Bangur CS, Lodes MJ, Fanger GR, Vedvick TS, Carter D;
PI Retter MW, Mannion J;
XX
XX WPI; 2001-071488/08.
XX
XX Lung tumor-associated proteins and the nucleic acids that encode them,
PT useful for preventing, diagnosing and treating lung cancer -
XX
XX Claim 4; Page 413; 436pp; English.
XX
XX The present invention describes immunogenic portions of lung tumour-
CC associated proteins (I) and the nucleic acids (NAs) that encode them.
CC (I) have cytostatic activity and can be used in gene therapy, antisense
CC inhibition and in vaccines. The NAs and the lung tumour-associated
CC proteins they encode may be used in the prevention, treatment and
CC diagnosis of diseases associated with their inappropriate expression,
CC especially lung cancers. For example, the NAs may be administered to
CC treat diseases by rectifying mutations or deletions in a patient's genome
CC that affect the activity of the protein by expressing inactive proteins
CC or to supplement the patients own production of (I). Additionally, the
CC NAs may be used to produce the lung-tumour associated protein, according
CC to standard recombinant DNA methodology. Conversely, antisense NA

CC molecules may be administered to down regulate protein expression by
CC binding with the cells own genes and preventing their expression. The NA
CC and complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar NA sequences in
CC samples, and hence which patients may be in need of treatment for lung
CC cancer. The (I) may be used as antigens in the production of antibodies
CC and in assays to identify modulators (agonists and antagonists) of the
CC expression and activity of the protein. AAF68083 to AAF68878 and
CC AAF76848 to AAF76878 represent human lung tumour protein related
CC nucleotide and protein sequences which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 461 BP; 150 A; 104 C; 123 G; 84 T; 0 other;
Query Match 44.8%; Score 350; DB 22; Length 461;
Best Local Similarity 94.8%; Pred No. 9; le-95;
Matches 379; Conservative 0; Mismatches 5; Indels 16; Gaps 1;
QY 383 GGCATGAGCGACACACACAAACACAGAACCCAGCCAGTCCCGAGGCCAGTAATGG 442
DB 62 GACACGCGGCACACACAAACACAGAACCCAGTCCCGAGGCCAGTAATGG 121
QY 443 AGAGCCCCAAAAGAAAGAACCCAGCAGCTGAAAAGTCGGGATCTACACCTGGCGCAGCAC 502
DB 122 AGAGCCCCAAAAGAAAGAACCCAGCAGCTGAAAAGTCGGGATCTACACCTGGCGCAGCAC 181
QY 503 AGAAGAAGATCAGGATACAGCTGAGATCCCAG-----TGCGCGCACATGG 546
DB 182 AGAAGAAGATCAGGATACAGCTGAGATCCCAGTCTGGGAAGGAAATGCGCGCATGG 241
QY 547 AAGGTGATCTGCAAGAGCTGCATCAGTCAACACATCGGGGATAAATCTGGATTGGTTCC 606
DB 242 AAGGTGATCTGCAAGAGCTGCATCAGTCAACACATCGGGGATAAATCTGGATTGGTTCC 301
QY 607 GCGTCAAGGTCAAGATAATACCTAAAGAGGAACACTGTAAATGCCAGAGCGGTGAA 666
DB 302 GCGTCAAGGTCAAGATAATACCTAAAGAGGAACACTGTAAATGCCAGAGCGGTGAA 361
QY 667 GAGCAACCAAGTTTAAATGAAGACAGCTGAAACACAGCGCAAGCTGTTTATATTTAG 726
DB 362 GAGCAACCAAGTTTAAATGAAGACAGCTGAAACACAGCGCAAGCTGTTTATATTTAG 421
QY 727 TATTGACTTAAACTATCTCAATAAAGTTTTCAGCTTTC 766
DB 422 TATTGACTTAAACTATCTCAATAAAGTTTTCAGCTTTC 461
RESULT 10
AAS38132
ID AAS38132 standard; cDNA; 391 BP.
XX
AC AAS38132;
XX
DT 17-DEC-2001 (first entry)
XX
DE Novel human diagnostic and therapeutic gene #1190.
XX
XX Human; cancer; breast; lung; colon; prostate; cytostatic; diagnostic; ss.
XX
XX Homo sapiens.
OS
XX WO200166753-A2.
PN
XX 13-SEP-2001.
PD
XX 09-MAR-2001; 2001WO-US07787.
PF
XX 09-MAR-2000; 2000US-0188609.
PR
XX (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;

QY 489 CCTGGCAGCACAGCAAGATCAGGATACAGCTGAGATCCAGTCCGCGACATGGA 548
Db 381 CTGTATCAAGAGAGAGATATCAGGGTGCGCTGAGATTCAGTGCCTGACCTGGAA 440
QY 549 GGTGATCTGCAAGAGCTGCATCAGTCAACACCGGGGATAAATCTGGATTTGGGTTCCGG 608
Db 441 GCGGATCTCCAGGAGTATGTCCAGCAAGAGCTGGGATGGATGTGAAGGT-GGTACTGA 499
QY 609 CTTCAAGGTGAAGATAATACCTTAAGAGAGAACACACTGTAAATGCCAGAGCAGTGAAGA 668
Db 500 TGTCAAGGGGAAGATTTCTACCAAAAGCAGAGCACCTTTAAATGCCAGAGCAGGTGAAGG 559
QY 669 GCAACACCAAGTTTAAATGAAGACAAGCTGAAACCAACGCAAGCTGGTTTATATTAGATA 728
Db 560 GAAATCACAGGTTTAAAGGAAGATAAGCTGAAACAACACAAACTCTTTTATATTAGATA 619
QY 729 TTTGAC-TTAACTATCTCAATAAAGTTTTCGAGCTTTCACCAACAAAAA 781
Db 620 TTTTACTTTTAAAAATATCTTAATAAAGTTTAAAGCTTTTTCACCAAAAAA 673

RESULT 14
AAD14983
ID AAD14983 standard; DNA; 611 BP.
XX
AC AAD14983;
XX
DT 01-NOV-2001 (first entry)
XX
DE Human NOV4 DNA.
XX
KW Human; NOVX; G-antigen; GAGE-like protein; interferon;
KW G-protein coupled receptor; GPCR; hepatocyte nuclear factor;
KW mast cell protease; gene therapy; proliferative disorder; cancer;
KW immune disorder; hepatic disorder; cirrhosis; viral infection;
KW hepatitis; neuroofactory system-related disorder; neurological disorder;
KW Parkinson's disease; infertility; autoimmune disease; arthritis;
KW multiple sclerosis; allergy; wound healing; cytostatic; neutropenic;
KW immunosuppressive; neuroprotective; vulnerary; hepatotropic; ds.
XX Homo sapiens.

OS
FH Key
FH 5'UTR Location/Qualifiers
FT 1..173
FT /*tag= a
FT CDS 174..521
FT /*tag= b
FT /*product= "Human NOV4 protein"
FT 522..611
FT /*tag= c
PN WO200161009-A2.
XX
PD 23-AUG-2001.
XX
PF 15-FEB-2001; 2001WO-US04828.
XX
PR 15-FEB-2000; 2000US-0182723.
PR 15-FEB-2000; 2000US-0182724.
PR 15-FEB-2000; 2000US-0182733.
PR 22-FEB-2000; 2000US-0183896.
PR 23-FEB-2000; 2000US-0184275.
PR 23-FEB-2000; 2000US-0184482.
PR 23-FEB-2000; 2000US-0184497.
PR 24-FEB-2000; 2000US-0184744.
PR 13-APR-2000; 2000US-0197083.
PR 10-AUG-2000; 2000US-0224157.
PR 18-SEP-2000; 2000US-0233405.
PR 27-SEP-2000; 2000US-0236060.
PR 02-JAN-2001; 2001US-0259414.
PR 18-JAN-2001; 2001US-0262454.
PR 14-FEB-2001; 2001US-0783429.

XX (CURA-) CURAGEN CORP.
PA Malyankar UM, Tchernev VT, Padigaru M, Taupier RJ, Spytek KA;
PI Majumder K, Guo X, Spaderna SK, Boldog FL;
XX WPI: 2001-514775/56.
DR P-PSDB: AAE08583.
XX Isolated novel polypeptides useful for diagnosis of and treating
PT cancer, infertility, autoimmune diseases, arthritis, multiple
PT sclerosis, allergies, wound healing and hepatic disorders -
XX Claim 9; Page 14; 140pp; English.
XX The pres-nt sequence is a human NOV4 DNA. The NOVX protein has homology
CC with one of g-antigen (GAGE)-like protein, interferon, G-protein coupled
CC receptor (GPCR), hepatocyte nuclear factor or mast cell protease. The
CC NOVX is useful for treating or preventing a pathology associated with
CC NOVX. It is also useful for determining the presence or amount of NOVX
CC DNA in a sample, for identifying a potential therapeutic agent and in
CC gene therapy. It is also useful for determining the presence of or
CC predisposition to a disease associated with altered levels of NOVX. It is
CC also useful for the diagnosis and treatment of proliferative disorders,
CC e.g., cancer, immune disorders, hepatic disorders, e.g., cirrhosis, viral
CC infections, e.g., hepatitis, neuroofactory system-related disorders,
CC neurological disorders, e.g., Parkinson's disease, infertility,
CC autoimmune diseases, arthritis, multiple sclerosis, allergies and wound
XX healing.
XX Sequence 611 BP; 196 A; 127 C; 152 G; 135 T; 1 other;
SQ

Query Match 25.8%; Score 201.8; DB 22; Length 611;
Best Local Similarity 78.4%; Pred. No. 3e-50;
Matches 279; Conservative 0; Mismatches 73; Indels 4; Gaps 3;
QY 429 GAGCCCACTATGAGAGCCGCCCAAAAGAACAGACAGCAGCTCAAGTCGGATCCTACA 488
Db 255 GAACCCACTGATGAAGAGCCCTAAAGAGAGAAACACCCCACTTAAAGTCGGAATCCTACA 314
QY 489 CTGGG--CAGCAGACAGAAGAAGATCAGGATACAGCTGAGATCCAGTCGCCGACATGG 546
Db 315 CCTGACTCNAGAAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAGTCGACCTGG 374
QY 547 AAGTGTATCTCGAGAGCTGCATCAGTCAACACCGGGGATAAATCTGGATTTGGTTC 606
Db 375 AAGCGATCTCCAGGAGCTATGTCAGACAAGACACTGGGGATGTGAAGGT-GGTACT 433
QY 607 GCGTCAAGGTGAAGATAATACCTAAAGAGGAACACTGTAAAATGCCAGACAGCTGAA 666
Db 434 GATGTCAAGGGGAGAGATTCTACCAAAAGCAGAGCACTTTAAATGCCAGAGCAGTGAA 493
QY 667 GAGCAACCAAGTTTAAATGAAGACAAGCTGAAACAACGCAAGCTGCTTTTATATTAGA 726
Db 494 GGGAAATCACAGGTTTAAAGGAAGATAAGCTGAAACAACACAAACTGTTTTATATTAGA 553
QY 727 TATTGAC-TTAACTATCTCAATAAAGTTTTCAGCTTTTCACCAAAAAA 781
Db 554 TATTTTACTTTTAAAAATATCTTAATAAAGTTTAAAGCTTTTCTCCAAAAA 609

RESULT 15
AAS69484
ID AAS69484 standard; cDNA; 503 BP.
XX
AC AAS69484;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #5288.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

Qy 405 ACAGAACACACAGCCAGTCCCGAGGAGCC 433
Db 180 ACAGAACACACAGCCAGTCCCGAGGAGCC 208

RESULT 17

AAI60530
ID AAI60530 standard; cdna; 532 BP.

XX AC AAI60530;

XX DT 22-OCT-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 4519.

XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia; ss.

XX OS Homo sapiens.

XX PN WO200153312-A1.

XX PD 26-JUL-2001.

XX PF 26-DEC-2000; 2000WO-US34263.

XX PR 21-JAN-2000; 2000US-0488725.

XX PR 25-APR-2000; 2000US-0552317.

XX PR 09-JUL-2000; 2000US-0598042.

XX PR 19-JUL-2000; 2000US-0620312.

XX PR 03-AUG-2000; 2000US-0653450.

XX PR 14-SEP-2000; 2000US-0662191.

XX PR 19-OCT-2000; 2000US-0693036.

XX PR 29-NOV-2000; 2000US-0727344.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

XX PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX DR WPI: 2001-442253/47.

XX DR P-PSDB; AAM41374.

XX PT Novel nucleic acids and polypeptides, useful for treating disorders

XX PT such as central nervous system injuries -

XX PS Claim 1; SEQ ID NO 4519; 10078pp; English.

XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and

XX CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,

XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful

XX CC in gene therapy. A composition containing a polypeptide or polynucleotide

XX CC of the invention may be used to treat diseases of the peripheral nervous

XX CC system, such as peripheral nervous injuries, peripheral neuropathy and

XX CC localised neuropathies and central nervous system diseases, such as

XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

XX CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

XX CC utilisation of the activities such as: immune system suppression,

XX CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

XX CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

XX CC assays for receptor activity, arthritis and inflammation, leukaemias and

XX CC C.N.S disorders.

XX CC Note: The sequence data for this patent did not form part of the printed

XX CC specification.

XX SQ Sequence 532 BP; 154 A; 108 C; 150 G; 120 T; 0 other;

Query Match 25.0%; Score 195.6; DB 22; Length 532;

Best Local Similarity 79.4%; Pred.No. 2e-48; Indels 4; Gaps 4;

Matches 281; Conservative 0; Mismatches 69; Indels 4; Gaps 4;

Qy 426 CAGGAGCCAGTAATGGAGAGCCCAAAAGAAAGAACAGCAGCTGAAAGTCGGGATCCT 485

Db 180 CTGGAGCCGGTGATGAGGAGCCTCAGCAAGAGGAACCACTGAAAGTCGGGATCCT 239

Qy 486 ACACCTGGCCAGCAGACAGAGAAGATCAGGATACAGCTGAGATCCCATGCCGACATG 545

Db 240 GCACCTGGTCCAGGAGAGAGAAGATCAGGCTGAGCTGAGACTCAAGTCCCTCACCTG 299

Qy 546 GAAAGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAAATCTGGATTTGGTTC 605

Db 300 GAAGCTGATCTCCAGGAGCTGTCTCAGTCAAGAGCTGGGATGAATCGGAGAT-GGTCC 358

Qy 606 CGCGCTCAAGGTGAAGATTAATACCTAAAGAGGAACACTGTGTAATGCCAGAA-CAAGGTG 664

Db 359 TGATGTCCAGGGAAGATTTCTGACAAAGTCAGAGCAATTTAAATGCCAGAGGAGGTG 418

Qy 665 AAGAGCAACCAAGTTTAAATCAAGACAGCTGAAACAAAGC-AAGCTGGTTTTATATT 723

Db 419 ACAGGCAACCCAGGTTTAAATGAAGACAGCTGAAACAAACAAACTGTTTTATCTA 478

Qy 724 AGATATTTGACTT-AAACTATCTCAATAAAGTTTTTCAGCTTTCACCAAAAAA 776

Db 479 AGATATTTGACTTAAAAATATCGAAATAAACTTTTTCAGCTTTCCTCCGAAAAA 532

RESULT 18

AAA42613/c

ID AAA42613 standard; cdna; 201 BP.

XX AC AAA42613;

XX DT 21-AUG-2000 (first entry)

XX DE Human secreted expressed sequence tag SEQ ID NO:1353.

XX KW Human; mouse; xenopus; rat; secreted expressed sequence tag; SEST;

XX KW expressed sequence tag; EST; probe; chemotactic; proliferative;

XX KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;

XX KW thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;

XX KW antiviral; antidiabetic; antisthmatic; vulnerable; antiparkinsonian;

XX KW antitumor; osteoprotective; neuroprotective; nootropic; antipsoriatic;

XX KW cerebroprotective; anticonvulsant; antidepressant; gene therapy;

XX KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;

XX KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;

XX KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;

XX KW central nervous system disorder; Alzheimer's disease; stroke;

XX KW Parkinson's disease; Huntington's disease; coagulation disorder;

XX KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;

XX KW tumour; infection; depression; psoriasis; ss.

XX OS Homo sapiens.

XX PN WO200021990-A1.

XX PD 20-APR-2000.

XX PF 15-OCT-1999; 99WO-US24205.

XX PR 15-OCT-1998; 98US-0104435.

XX (GEMY) GENETICS INST INC.

XX PI Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;

XX PI Merberg D, Treacy M;

XX DR WPI: 2000-317937/27.

XX PT Isolated polynucleotides, and encoded proteins, comprising secreted

XX PT expressed sequence tags (seSTs), useful for treating various disorders

PT such as autoimmune, infectious, and central nervous system disorders -
XX Claim 1; Page 447; 618pp; English.
XX
XX AAA41261 to AAA43419 represent specifically claimed secreted expressed
CC sequence tags (sESTs), isolated from human, mouse, xenopus and rat
CC tissue sources. The sESTs can have a range of activities depending on
CC the tissues they were isolated from. The activities include:
CC chemotactic; proliferative; immunomodulatory; haematopoietic;
CC chemokine; analgesic; haemostatic; thrombolytic; antiinflammatory;
CC cytosatic; antibacterial; antifungal; antitumor; antidiabetic;
CC antiasthmatic; vulnary; antitumor; osteopathic; neuroprotective;
CC neurotropic; antiparkinsonian; antipsoriatic; cerebroprotective;
CC anticonvulsant; and antidepressant. The sESTs can be used for gene
CC therapy and in vaccines. The sESTs are useful as probes for the
CC identification and isolation of full-length cDNAs and genomic DNA
CC molecules which correspond to the sESTs. Proteins encoded by the sESTs
CC are useful in assays for determining biological activity and raising
CC antibodies. They may be useful for treatment of autoimmune disorders
CC (multiple sclerosis, insulin dependent diabetes), allergic conditions
CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,
CC osteoporosis, osteoarthritis, central nervous system disorders
CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
CC disease), tumours, bacterial, fungal or viral infections, depression and
CC psoriasis. AAA43420 to AAA43425 represent linker variants which are given
CC in the exemplification of the present invention.
XX
XX Sequence 201 BP; 33 A; 56 C; 42 G; 70 T; 0 other;

Query Match 22.4%; Score 174.8; DB 21; Length 201;
Best Local Similarity 98.9%; Pred. No. 2.1e-42;
Matches 176; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 534 GTGCGGCACATGGAGGTGATCTCCAGAGCTGCATCAGTCAACACCGGGGATAATCT 593
Db 198 GAGAGCGACATGGAGGTGATCTCCAGAGCTGCATCAGTCAACACCGGGGATAATCT 139
Qy 594 GGATTTGGTTCGCGGTCAAGGTGAAGATAAATACCTAAAGAGGAACACTGTAAATGCC 653
Db 138 GGATTTGGTTCGCGGTCAAGGTGAAGATAAATACCTAAAGAGGAACACTGTAAATGCC 79
Qy 654 AGAAGCAGGTGAAGAGCAACCAACCAAGTTTAAATGAAGACAGCTGAACACGCAAGC 711
Db 78 AGAAGCAGGTGAAGAGCAACCAACCAAGTTTAAATGAAGACAGCTGAACACGCAAGC 21

RESULT 19
AAS69486
ID AAS69486 standard; cDNA; 661 BP.
XX
XX AAS69486;
AC
XX
XX 13-FEB-2002 (first entry)
DT
XX
XX DNA encoding novel human diagnostic protein #5290.
DE
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
KW
XX
XX Homo sapiens.
OS
XX
XX WO200175067-A2.
PN
XX
XX 11-OCT-2001.
PD
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
PR
XX
XX 23-AUG-2000; 2000US-0649167.
PR
XX
XX (HYSE-) HYSEQ INC.

PA
XX

PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR P-PSDB; ABG05299.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX Claim 1; SEQ ID No 5290; 103pp; English.
PS
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 661 BP; 212 A; 129 C; 158 G; 162 T; 0 other;

Query Match 22.0%; Score 171.8; DB 23; Length 661;
Best Local Similarity 76.9%; Pred. No. 3.2e-41;
Matches 273; Conservative 0; Mismatches 77; Indels 5; Gaps 5;

Qy 426 CAGGAGCCCATGATGAGAGCCCAAAAAGAACAGAACAGCAGCTGAAAGTCGGATCCT 485
Db 307 CTGAGCCCGGTGATGAGAGCCCTCAGCAAGAGGAACCACTGAAAGTCGGATCCT 366
Qy 486 ACACCTGGG-CAGCAGACAGAGAAGATCAGGATACAGCTGAGATCCCGCAGCAT 544
Db 367 GCACCTGGGTGAGAGAGAGAAGATCAGGTCAGCTGAGACTCAAGTCCTGACCT 426
Qy 545 GGAAGG-TGATCTGCAAGAGCTGCATCAGTCAACACCGGGGATAAATCTGGATTGGGT 603
Db 427 GGAAGGCTGATCTCCAGGAGCTGTCTCAGTCAAGAGCTGGGGGTGAATGTGGAATGGTC 486
Qy 604 TCCGGCGTCAAGGTGAAGATAATACCTAAAGAGGAACACTGTAAATGCCAGAA-GCAGG 662
Db 487 CTGATGACCAGGGGAGAGATTCTGCCAATAATCAGAACAAATTTTAAATGCCAGAGGAGG 546
Qy 663 TGAAGAGCAACCAAGTTTAAATGAAGACAGCTGAACCAACGC-AAGCTGGTTTATA 721
Db 547 TGACAGGCAACCAAGTTTAAATGAAGACAGCTGAACCAACCAACCAACTGTTTTATC 606
Qy 722 TTAGATATTTGACTT-AAACTATCTCAATAAAGTTTTCAGCTTTCACCAAAAAA 775
Db 607 TAAGATATTTGACTTAAATAATATCGAATAAATCTTTCAGCTTTCCTCAAAAAA 661

RESULT 20
AAF59637
ID AAF59637 standard; cDNA; 580 BP.
XX
XX AAF59637;
AC
XX
XX 24-APR-2001 (first entry)
DT
XX

XX P1 Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR P-PSDB; ABG27048.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX Claim 1: SEQ ID NO 27039; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 750 BP; 225 A; 164 C; 203 G; 158 T; 0 other;
SQ

Query Match 17.8%; Score 139.2; DB 23; Length 750;
Best Local Similarity 77.1%; Pred. No. 2e-31;
Matches 182; Conservative 0; Mismatches 53; Indels 1; Gaps 1;
QY 426 CAGGAGCCAGTAATGAGAGCCCAAAAGAGAACAGCAGCTGAAAGTCGGGATCCT 485
DB 208 CAGGAGCCAGTGTCCAGAGCCTCACNAGAGAACCAACCACTGAAAGTCAGGATCAT 267
QY 486 ACACCTGGCGCAGCAGACAGAGAAGATCAGGATACAGCTGAGATCCAGTCGCCGACATG 545
DB 268 ACACCTGGTCAGAGAGAGAAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTAACCTG 327
QY 546 GAAGGTGATCTGCAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTGGGTTC 605
DB 328 GAAGCTGATCTCCAGGAGCTGTCTCAGTCAAAAGACTGGGGATGAATGCGGAGAT-AGTCC 386
QY 606 CGCGCTCAAGGTGAAGATAATACCTAAAGAGGAACACCTGTAAATGCCAGAGACAG 661
DB 387 TGATGTCCAGGGAAGATTCTCCAAAATCAGAGCAATTTAAATATGCAGAGGAG 442

RESULT 24
AAH83260
ID AAH83260 standard; cDNA; 320 BP.
AC
XX
XX AAH83260;
XX
XX 25-SEP-2001 (first entry)
XX
XX Human ovarian tumour associated polynucleotide sequence SEQ ID NO:884.
XX
XX Human; ovarian tumour; ovarian cancer; diagnosis; gene therapy;
KW immunogenic; vaccine; ss.
XX
XX Homo sapiens.
OS

PN WO200151513-A2.
XX
XX 19-JUL-2001.
PD
XX
XX 16-JAN-2001; 2001WO-US01575.
PF
XX
XX 14-JAN-2000; 2000US-0176722.
PR
XX
XX (CORI-) CORIXA CORP.
PA
XX
XX Algate PA;
PI
XX
XX WPI; 2001-425866/45.
DR
XX
XX Novel ovarian tumor proteins, and nucleic acids encoding them, used to
PT treat and diagnose cancers, particularly ovarian cancer -
PT
XX
XX Claim 5; Page 225; 338pp; English.
PS
XX
XX AAH82377 to AAH83878 represent human ovarian tumour-associated
CC polynucleotide sequences which encode ovarian tumour proteins. The
CC ovarian tumour protein and polynucleotide sequences have cytostatic
CC activity, and can be used in gene therapy and vaccine production. The
CC ovarian tumour proteins and polynucleotides can be used to inhibit
CC the development of cancer, particularly ovarian cancer. They can also
CC be used to diagnose the onset and progression of cancer.
XX
XX Sequence 320 BP; 96 A; 65 C; 85 G; 74 T; 0 other;
SQ

Query Match 17.3%; Score 135; DB 22; Length 320;
Best Local Similarity 79.0%; Pred. No. 2.4e-30;
Matches 173; Conservative 0; Mismatches 45; Indels 1; Gaps 1;
QY 523 CTGAGATCCCATGCGGCACATGGAAGTGTCTGCAAGAGCTGCATCAGTCAAAACCG 582
DB 1 CTGAGATTCAAGTGCCTGACCTGGAGCCGATCTCCAGGAGCTATGTCAGACAAGACTG 60
QY 583 GGGATAAATCTGGATTGGGTTCCGGCTCAAGGTGAAGATAATACCTAAAGAGAACAC 642
DB 61 GGGATGGATGTGAAGGT-GGTACTGATGCAAGGGAAGATTCTACCAAAAAGCAGAGCAC 119
QY 643 TGTAAATGCCAGAGCAGGTGAAGAGCAACCAACCAAGTTTAAATCAACAGCAAGCTGAAC 702
DB 120 TTTAAATGCCAGAGCAGGTGAAGGGAATACACAGGTTTAAAGGAAGATAAGCTGAAC 179
QY 703 AACGCAAGCTGGTTTATATTAGATATTGACTTAAACT 741
DB 180 AACACAAACTGTTTTATATTAGATATTTTACTTTACCT 218

RESULT 25
AAS24637
ID AAS24637 standard; cDNA; 214 BP.
XX
XX AAS24637;
AC
XX
XX 07-NOV-2001 (first entry)
DT
XX
XX Human ovarian PCR-subtracted cDNA library clone #818.
DE
XX
XX Immunogenic protein; cancer; ovarian tumour; T-cell stimulation; ss;
KW gene therapy; cytostatic; T-cell expansion; nucleic acid hybridisation;
KW primer; probe.
XX
XX Homo sapiens.
OS
XX
XX WO200157207-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 05-FEB-2001; 2001WO-US03733.
PF
XX
XX 04-FEB-2000; 2000US-0180403.
PR

```
PR 28-MAR-2000; 2000US-0192745.
PA (CORI-) CORIXA CORP.
PI Algate PA, Mannion J;
XX WPI: 2001-488879/53.
XX
XX New polynucleotides encoding ovarian tumour proteins, useful for
PT treating ovarian cancer, and as probes, primers, and markers of cancer
PT progression
XX
XX Example 1: page 241-242; 378pp; English.
XX
XX The invention comprises compositions used for the therapy and diagnosis
CC of ovarian cancer. The compositions comprise one or more ovarian tumour
CC proteins, their associated polynucleotides, or immunogenic portions of
CC the proteins. The ovarian tumour polynucleotides and polypeptides are
CC useful for stimulating and/or expanding T cells specific for a tumour
CC protein. They are also useful for inhibiting the development of cancer in
CC a patient with an ovarian tumour DNA or protein by incubating isolated
CC T-cells allowing them to proliferate, and administering to the patient.
CC The sequences can be used as markers for cancer, for example, to monitor
CC ovarian cancer progression. Probes and primers are useful in nucleic acid
CC hybridisation, in detecting the presence of complementary sequences in a
CC given sample, for preparing mutant species and for preparing other
CC genetic constructions. Sequences AAS23820-AAS25231 and AAS25328-AAS25549
CC represent human ovarian tumour protein cDNA clones.
XX
XX Sequence 214 BP; 74 A; 34 C; 55 G; 51 T; 0 other;
SQ
Query Match 17.1%; Score 133.2; DB 22; Length 214;
Best Local Similarity 79.4%; Pred. No. 6.7e-30;
Matches 170; Conservative 0; Mismatches 43; Indels 1; Gaps 1;
QY 523 CTGAGATCCCATCGGCGACATGAAGTGATCTGCAAGAGCTGCATCAGTCAACACCG 582
Db 1 CTGAGATTCAGTGCCTGACCTGGAACCGCATCTCCAGGAGCTATGTACACAAGACTG 60
QY 583 GGGATAAATCTGGATTTGGGTTCCTGGGCTCAAGTGAAGATAATACCTAAAGAGGAAC 642
Db 61 GGGATGGATGTGAAGGT-GGTACTGATGTCGAAGGGGAAGATCTTACCAAAAGCAGAC 119
QY 643 TCTAAATGCCAGAGCAGGTGAAGAGCAACCAAGCTTTAAATGAAGACAAGCTGAAC 702
Db 120 TTTAAATGCCAGAGCAGGTGAAGGGAAATCACAGGTTTAAAGGAAGATAAGCTGAAC 179
QY 703 AACGCAAGCTGGTTTATATTAGATATTGACTT 736
Db 180 AACACAACACTGTTTTATATTAGATATTTTACTT 213
RESULT 26
AAC59112/c
ID AAC59112 standard; cDNA; 2182 BP.
XX
XX AAC59112;
AC
XX
XX 02-FEB-2001 (first entry)
DE
XX Human secreted protein coding sequence SEQ ID NO: 15.
XX
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200055177-A2.
PN
XX
XX 21-SEP-2000.
PD
```

```
XX 09-MAR-2000; 2000WO-US06058.
XX
XX 12-MAR-1999; 99US-0124145.
PR 03-DEC-1999; 99US-0168654.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Komatsoulis G;
PI
XX WPI: 2000-638177/61.
DR P-PSDB; AAB28016.
XX
XX Novel nucleic acids encoding 49 human secreted proteins useful for
PT treating cancers, hyperproliferative disorders, inflammatory disorders,
PT neurological disorders and cardiovascular disorders
XX
XX Claim 1: Page 319-320; 389pp; English.
XX
XX The invention relates to the isolation of genes AAS9108-A59156 encoding
CC 49 human secreted proteins AAB28012-B28060. The genes can be used to
CC generate fusion proteins by linking to the gene for the human
CC immunoglobulin G Fc portion (SEQID) for increasing the stability of
CC the fusion protein as compared to the human protein only. The genes and
CC proteins are useful for preventing, ameliorating or treating medical
CC conditions, e.g. by protein or gene therapy. The genes are isolated
CC from a range of human tissues disclosed in the specification. The
CC nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer, and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d)
CC wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
XX
XX Sequence 2182 BP; 502 A; 639 C; 579 G; 462 T; 0 other;
SQ
Query Match 17.0%; Score 132.6; DB 21; Length 2182;
Best Local Similarity 78.5%; Pred. No. 3.4e-29;
Matches 172; Conservative 0; Mismatches 44; Indels 3; Gaps 1;
QY 14 GAGCCCGCGACTCGGTCCTGAGGTCTGGATCTTTTTCGGTACTGAGACAGCGGG 73
Db 231 GAGCTGTGAGGTCAAGTCTCTGAGGTCTGGATTATTTCTCTCTACTGAGACGAGCAG 172
QY 74 TAGGTCCACAGGCAGATCCAACTGGGAGTTCAGTGTGAGTGAAGAGTGAACAGCAAC 133
Db 171 TAGGTCCACAGGCAGATCCAACTGGGAGTTCAGTGTGAGTGAAGAGTGAAGAGGAGCCAG 112
QY 134 CAGGCTTCCCGAGGGTTGT---GTGGTCAGTGAAGTGAAGAGCCCTCGAAGTCG 190
Db 111 CAGGCTTCCAGAGGGTTCATCAGGGGGAGGGAGACTCAGAGGGGAGAGGAGGAGGCTCT 52
QY 191 TCGTCCCTCTCATCGCGGTGCCACGCCCATGGACCTTCTT 229
Db 51 TCATCCTCTTTCATGCGCCACAGCAGCGGGGCTCTT 13
RESULT 27
AAV18721
ID AAV18721 standard; cDNA; 539 BP.
XX
XX AAV18721;
AC
XX
XX 30-JUL-1998 (first entry)
DE
XX cDNA encoding GAGE-6 tumour rejection antigen precursor.
XX
XX GAGE tumour rejection antigen precursor; TRAP; tumour;
```

diagnosis; melanoma; antigen; cytolytic T cell clone proliferation;
HLA-typing assay; ss.

Homo sapiens.

Key Location/Qualifiers
CDS 82..436
/*tag= a

FT /transl_except= (pos:127..129, aa:Arg)
FT /transl_except= (pos:196..198, aa:Ala)
FT /transl_except= (pos:199..201, aa:Thr)

PN WO9749417-A1.

XX 31-DEC-1997.

XX 23-JUN-1997; 97WO-US10850.

XX 24-JUN-1996; 96US-0669161.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Boon-Falleur T, Debacker O, Van Den Eynde B;

XX WPI; 1998-076905/07.

XX P-PSDB; AAW47603.

XX Isolated nucleic acid encoding GAGE tumour rejection antigen

PT precursor - processed by HLA-Cw6 molecules into peptides, useful to

PT diagnose melanomas

XX Example 13; Fig 4; 60pp; English.

XX The present sequence encodes a GAGE-6 tumour tumour rejection antigen
CC precursor (TRAP). The protein is expressed in a number of tumours. In
CC contrast the only normal tissue which expresses GAGE TRAP protein is
CC testis. Several GAGE TRAPs have been identified (see AAV18717-21). The
CC major difference between these proteins and GAGE-1 (AAV05540) is the
CC absence of a stretch of 143 bases located at position 379 to 521 of the
CC GAGE-1 TRAP sequence. The rest of the sequences show mismatches at
CC various position, with the exception of GAGE-3 whose 5' end is totally
CC different from the other GAGE cDNAs for the first 112 bases. This
CC region of GAGE-3 cDNA contains a long repeat and a hairpin structure.
CC The antigens can be used to diagnose melanomas, characterised by
CC expression of a TRAP or presentation of a tumour rejection antigen.
CC Antigens shed into blood or urine can be observed and then used to
CC confirm a diagnosis of melanoma using cytolytic T cell clone
CC proliferation methodologies. Other uses for the processed peptides,
CC include HLA-typing assays for, e.g. skin graft or organ transplants.

XX Sequence 539 BP; 157 A; 114 C; 157 G; 111 T; 0 other;

Query Match 14.3%; Score 111.6; DB 19; Length 539;
Best Local Similarity 64.2%; Pred. No. 3.3e-23;
Matches 217; Conservative 0; Mismatches 114; Indels 7; Gaps 3;

QY 447 CCCCCAAAGAAGACACGAGCTGAAGTGGGATCTACACCTGGCGAGCAGACAGAA 506

DB 202 CCTGAAGAAGGGGAACACGACCACTCAAGCTCAGGATCTGCAGCTGCTCAGGAG---GGA 258

QY 507 GAAGATCAGGATACAGCTGAGATCCAGTGGCGGACATGGAAGTGATCTCAAGAGCTG 566

DB 259 GAGGATGAGGAGCATCTGCAGGTCAAGGGCCGAGCCTGAAGCTGATAGCCAGGAACAG 318

QY 567 CATCAGTCAAAACACCGGGGATAAATCTGGATTTGGTTCGGCGCTCAAGGTGAAGATAAT 626

DB 319 GGTCAACCCACACACTGGGTGTGAGTGTGAAGAT--GGTCTGTATGGCAGGAGGTGGACCC 377

QY 627 ACCTAAAGAGGAACACTGTAAATCCAGACGAGTGAAGAGCAACCAACCAAGTTTAAAT 686

DB 378 GCCAATCCAGGAGGTGAACACCCCTGAAGAAGGTGAAGAAGCAATCACAGTGTGTTAAA 437

QY 687 GAAGACAAGCTGAACAAACGCAAGCTGTTTATATTAGATATTG---ACTTAAACTAT 743

DB 438 GAAGACAGCTGAATATGATGAGGCTGCTCTATGTGGAATTTGTTCAITTAATAATCT 497

QY 744 CTCATAAAGTTTTCAGCTTTACCAAAAAA 781

DB 498 CCAAATAAGCTTTACAGCCTTCTCCAAAAA 535

RESULT 28

AAAX90521

ID AAAX90521 standard; cdna; 540 BP.

XX AC AAAX90521;

XX DT 30-SEP-1999 (first entry)

XX GAGE-4 tumour rejection antigen clone nucleotide sequence.

XX Human leukocyte antigen; HLA-A29; tumour rejection antigen;

XX detection; therapy; pathological condition; cancer; CTL;

XX cytolytic T lymphocyte; GAGE; ss.

XX OS Homo sapiens.

XX PN WO9937665-A1.

XX PD 29-JUL-1999.

XX PF 12-JAN-1999; 99WO-US00775.

XX PR 23-JAN-1998; 98US-0012818.

XX PA (LUDW-) LUDWIG INST CANCER RES.

XX PI Boon-Falleur T, Debacker O, Van Den Eynde B, Van Der Bruggen P;

XX DR WPI; 1999-469111/39.

XX PT New isolated peptides which bind to HLA-A29 molecules, which are

XX tumour rejection antigens used for detection and therapy of

XX pathological conditions, e.g. cancer

XX Example 13; Fig 4; 62pp; English.

XX The present invention describes peptides which bind to human leukocyte
CC antigen (HLA) A29 (HLA-A29) molecules. The peptides are processed into
CC tumour rejection antigens. They can be used for detecting cytolytic T
CC lymphocytes (CTLs) in pathological conditions such as cancer and in
CC HLA-typing assays. Complexes of HLA-A29 molecules and the peptides can
CC be used for stimulating CTLs in vivo. The present sequence represents
CC a GAGE tumour rejection antigen clone, from an example from the present
CC invention.

XX Sequence 540 BP; 159 A; 114 C; 156 G; 111 T; 0 other;

Query Match 14.1%; Score 110; DB 20; Length 540;
Best Local Similarity 63.9%; Pred. No. 9.9e-23;
Matches 216; Conservative 0; Mismatches 115; Indels 7; Gaps 3;

QY 447 CCCCCAAAGAAGACACGAGCTGAAGTGGGATCTACACCTGGCGACACACAGAA 506

DB 203 CCTGAAGAAGGGGAACACGACCACTCAAGCTCAGGATCTGCAGCTGCTCAGGAG---GGA 259

QY 507 GAAGATCAGGATACAGCTGAGATCCAGTGGCGGACATGGAAGTGATCTCAAGAGCTG 566

DB 260 GAGGATGAGGAGCATCTGCAGGTCAAGGGCCGAGCCTGAAGCTGATAGCCAGGAACAG 319

QY 567 CATCAGTCAAAACACCGGGGATAAATCTGGATTTGGTTCGGCGCTCAAGGTGAAGATAAT 626

DB 320 GGTCAACCCACAGCTGGGTGTGAGTGTGAAGAT--GGTCTGTATGGCAGGAGTGGACCC 378

QY 627 ACCTAAAGAGGAACACTGTAAATGCCAGAGCAGGTGAAGAGCAACCAACCAAGTTTAAAT 686

```
Db 379 GCCAATCCAGAGGAGGTGAARACGCTGAAGAAGCTGAAAGCAATCACAGTGTAAAA 438
Qy 687 GAGACAAAGCTGAACAACGCAAGCTGTTTATATAGATATTTG---ACTTAAACTAT 743
Db 439 GAAGGACGTTGAATCATGACAGGCTGCTCTATGTGGAAATTTGTTCAATTAATCT 498
Qy 744 CTCATAAAGCTTTTGAGCTTTCCACCAAAAAA 781
Db 499 CCAATAAAGCTTTACAGCCTTCTGCAAAAAA 536

RESULT 29
ID AAX90523 standard; cDNA; 540 BP.
XX
AC AAX90523;
XX
DT 30-SEP-1999 (first entry)
XX
DE GAGE-6 tumour rejection antigen clone nucleotide sequence.
XX
KW Human leukocyte antigen; HLA-A29; tumour rejection antigen;
KW detection; therapy; pathological condition; cancer; CTL;
KW cytolytic T lymphocyte; GAGE; ss.
XX
OS Homo sapiens.
XX
XX WO9937665-A1.
XX
PD 29-JUL-1999.
XX
XX 12-JAN-1999; 99WO-US00775.
XX
XX 23-JAN-1998; 98US-0012818.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Boon-Falleur T, Debacker O, Van Den Eynde B, Van Der Bruggen P;
XX WPI; 1999-469111/39.
XX
XX New isolated peptides which bind to HLA-A29 molecules, which are
XX tumour rejection antigens used for detection and therapy of
XX pathological conditions, e.g. cancer
XX
XX Example 13; Fig 4; 62pp; English.
XX
XX The present invention describes peptides which bind to human leukocyte
XX antigen (HLA) A29 (HLA-A29) molecules. The peptides are processed into
XX tumour rejection antigens. They can be used for detecting cytolytic T
XX lymphocytes (CTLs) in pathological conditions such as cancer and in
XX HLA-typing assays. Complexes of HLA-A29 molecules and the peptides can
XX be used for stimulating CTLs in vivo. The present sequence represents
XX a GAGE tumour rejection antigen clone, from an example from the present
XX invention.
XX
XX Sequence 540 BP; 159 A; 113 C; 157 G; 111 T; 0 other;
XX
XX Query Match 14.1%; Score 110; DB 20; Length 540;
XX Best Local Similarity 63.9%; Pred. NO. 9.9e-23;
XX Matches 216; Conservative 0; Mismatches 115; Indels 7; Gaps 3;

Qy 447 CCCCCAAAAGAACACAGCAGCTGAAAGTCGGGATCCTACACTGGCCAGCAGACAGAA 506
Db 202 CCTGAAGAGGGGAAACCACTCAACGTCAGGATCCTCGAGTGTCTAGGAG---GGA 258
Qy 507 GAAGATCAGGATCAGCTGAGATCCCAAGTCGGGACATGGAAGTGATCTGCAAGAGCTG 566
Db 259 GAGGATGAGGGAGCATCTGCAAGTCAAGGGCCGAAGCCTGAAGCTGATAGCCAGGAACAG 318
Qy 567 CATCAGTCAACACCGGGGATAAATCTGGATTGGGTTCCGGCTCAAGGTGAAGATAAT 626
Db 319 GGTCAACCCACAGACTGGGTGTGAGTGTGAAGAT-GGTCTGTGGGAGGAGGTGGACCC 377
```

```
Qy 627 ACCTAAGAGGAAACACTGTAAATGCCAAGACAGCTGAAGACCAACCAAGTTTAAAT 686
Db 378 GCCAATCCAGAGGAGCTGAACACGCTGAAGAAGCTGAAAGCAATCACAGTGTAAAA 437
Qy 687 GAGACAAAGCTGAACAACGCAAGCTGTTTATATAGATATTTG---ACTTAAACTAT 743
Db 438 GAAGACACGTTGAATCATGACAGGCTGCTCTATGTGGAAATTTGTTCAATTAATCT 497
Qy 744 CTCATAAAGCTTTTGAGCTTTCCACCAAAAAA 781
Db 498 CCAATAAAGCTTTACAGCCTTCTGCAAAAAA 535
```

```
RESULT 30
AAV18720
ID AAV18720 standard; cDNA; 532 BP.
XX
AC AAV18720;
XX
DT 30-JUL-1998 (first entry)
XX
DE cDNA encoding GAGE-5 tumour rejection antigen precursor.
XX
KW GAGE tumour rejection antigen precursor; TRAP; tumour;
KW diagnosis; melanoma; antigen; cytolytic T cell clone proliferation;
KW HLA-typing assay; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 75..429
XX FT /*tag= a
XX FT /transl_except= (pos:189..191, aa:Ala)
XX FT /transl_except= (pos:192..194, aa:Thr)
XX
XX WO9749417-A1.
XX
XX 31-DEC-1997.
XX
XX 23-JUN-1997; 97WO-US10850.
XX
XX 24-JUN-1996; 96US-0669161.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Boon-Falleur T, Debacker O, Van Den Eynde B;
XX WPI; 1998-076905/07.
XX P-PSDB; AAW47602.
XX
XX Isolated nucleic acid encoding GAGE tumour rejection antigen
XX precursor - processed by HLA-Cw6 molecules into peptides, useful to
XX diagnose melanomas
XX
XX Example 13; Fig 4; 60pp; English.
XX
XX The present sequence encodes a GAGE-5 tumour tumour rejection antigen
XX precursor (TRAP). The protein is expressed in a number of tumours. In
XX contrast the only normal tissue which expresses GAGE TRAP protein is
XX testis. Several GAGE TRAPs have been identified (see AAV18717-21). The
XX major difference between these proteins and GAGE-1 (AAV05540) is the
XX absence of a stretch of 143 bases located at position 379 to 521 of the
XX GAGE-1 TRAP sequence. The rest of the sequences show mismatches at
XX various position, with the exception of GAGE-3 whose 5' end is totally
XX different from the other GAGE cDNAs for the first 112 bases. This
XX region of GAGE-3 cDNA contains a long repeat and a hairpin structure.
XX The antigens can be used to diagnose melanomas, characterised by
XX expression of a TRAP or presentation of a tumour rejection antigen.
XX Antigens shed into blood or urine can be observed and then used to
XX confirm a diagnosis of melanoma using cytolytic T cell clone
XX proliferation methodologies. Other uses for the processed peptides,
XX include HLA-typing assays for, e.g. skin graft or organ transplants.
```

XX SQ Sequence 532 BP; 154 A; 110 C; 156 G; 112 T; 0 other;

Query Match 13.9%; Score 108.4; DB 19; Length 532;
 Best Local Similarity 63.6%; Pred. No. 3e-22;
 Matches 215; Conservative 0; Mismatches 116; Indels 7; Gaps 3;

QY 447 CCCCCAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCCTACACCTGGGCAGCAGACAGAA 506
 DB 195 CTTGAAGAAGGGGAACACAGCAACTCAAGCTCAGGATCCTGCAGCTGCTCAGGAG---GGA 251

QY 507 GAAGATCAGGATACAGCTGAGATCCCAGTGGCGACATGGAAGTGATCTGCAAGAGCTG 566
 DB 252 GAGGATGAGGAGCAGCTGCAAGTCAAGGCCGAAGCTGAAGCTGATAGCCAGGAACAG 311

QY 567 CATCAGTCAAAACACCGGGGATAAATCTGGATTGGGTTCGGCCTCAAGCTGAAGATAAT 626
 DB 312 GGTCAACCCACAGACTGGGTGTGAGTGTGAAGAT--GGTCCTGATGGCAGGAGATGGACCC 370

QY 627 ACCTAAGAGGAACACTGTAATAATGCCAGAACAGCTGAAGCAACACCAAGTTTAAAT 686
 DB 371 GCCAAATCCAGAGAGGTGAAACGCCCTGAAGAAGGTGAAAGCAATCACAGTGTAAAA 430

QY 687 GAAGCAAGCTGAACACAGCAAGCTGGTTTATATTAGATATTG---ACTTAACTAT 743
 DB 431 GAAGCAGCTGTAATGATGAGGCTGCTCTATGTTGGAATTTGTTCAATAAATCT 490

QY 744 CTCAATAAAGTTTGCAGCTTTCACCAAAAAA 781
 DB 491 CCCAATAAAGCTTTACAGCCTTCTGCAAGAAAAA 528

RESULT 31
 AAX90522
 ID AAX90522 standard; cDNA; 532 BP.
 AC AAX90522;
 XX
 XX 30-SEP-1999 (first entry)
 XX GAGE-5 tumour rejection antigen clone nucleotide sequence.
 DE Human leukocyte antigen; HLA-A29; tumour rejection antigen;
 KW detection; therapy; pathological condition; cancer; CTL;
 KW cytolytic T lymphocyte; GAGE; ss.
 XX Homo sapiens.
 OS
 XX WO9937665-A1.
 PN
 XX 29-JUL-1999.
 PD
 XX 12-JAN-1999; 99WO-US00775.
 PF
 XX 23-JAN-1998; 98US-0012818.
 PR
 XX (LUDW-) LUDWIG INST CANCER RES.
 PA
 XX Boon-Falleur T, Debacker O, Van Den Eynde B, Van Der Bruggen P;
 PI WPT; 1999-469111/39.
 XX
 XX New isolated peptides which bind to HLA-A29 molecules, which are
 PT tumour rejection antigens used for detection and therapy of
 PT pathological conditions, e.g. cancer
 PS
 XX Example 13; Fig 4; 62pp; English.
 XX
 CC The present invention describes peptides which bind to human leukocyte
 CC antigen (HLA) A29 (HLA-A29) molecules. The peptides are processed into
 CC tumour rejection antigens. They can be used for detecting cytolytic T
 CC lymphocytes (CTLs) in pathological conditions such as cancer and in
 CC HLA-typing assays. Complexes of HLA-29 molecules and the peptides can

CC be used for stimulating CTLs in vivo. The present sequence represents
 CC a GAGE tumour rejection antigen clone, from an example from the present
 CC invention.

XX SQ Sequence 532 BP; 156 A; 111 C; 154 G; 111 T; 0 other;

Query Match 13.9%; Score 108.4; DB 20; Length 532;
 Best Local Similarity 63.6%; Pred. No. 3e-22;
 Matches 215; Conservative 0; Mismatches 116; Indels 7; Gaps 3;

QY 447 CCCCCAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCCTACACCTGGGCAGCAGACAGAA 506
 DB 195 CTTGAAGAAGGGGAACACAGCAACTCAAGCTCAGGATCCTGCAGCTGCTCAGGAG---GGA 251

QY 507 GAAGATCAGGATACAGCTGAGATCCCAGTGGCGACATGGAAGTGATCTGCAAGAGCTG 566
 DB 252 GAGGATGAGGAGCAGCTGCAAGTCAAGGCCGAAGCTGAAGCTGATAGCCAGGAACAG 311

QY 567 CATCAGTCAAAACACCGGGGATAAATCTGGATTGGGTTCGGCCTCAAGCTGAAGATAAT 626
 DB 312 GGTCAACCCACAGACTGGGTGTGAGTGTGAAGAT--GGTCCTGATGGCAGGAGATGGACCC 370

QY 627 ACCTAAGAGGAACACTGTAATAATGCCAGAACAGCTGAAGCAACACCAAGTTTAAAT 686
 DB 371 GCCAAATCCAGAGAGGTGAAACGCCCTGAAGAAGGTGAAAGCAATCACAGTGTAAAA 430

QY 687 GAAGCAAGCTGAACACAGCAAGCTGGTTTATATTAGATATTG---ACTTAACTAT 743
 DB 431 GAAGCAGCTGTAATGATGAGGCTGCTCTATGTTGGAATTTGTTCAATAAATCT 490

QY 744 CTCAATAAAGTTTGCAGCTTTCACCAAAAAA 781
 DB 491 CCCAATAAAGCTTTACAGCCTTCTGCAAGAAAAA 528

RESULT 32
 AAV18717
 ID AAV18717 standard; cDNA; 535 BP.
 AC AAV18717;
 XX
 XX 30-JUL-1998 (first entry)
 XX
 DE cDNA encoding GAGE-2 tumour rejection antigen precursor.
 XX GAGE tumour rejection antigen precursor; TRAP; tumour;
 KW diagnosis; melanoma; antigen; cytolytic T cell clone proliferation;
 KW HLA-typing assay; ss.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH 81..431
 FT CDS /*tag= a
 FT /transl_except= (pos:192..194, aa:Ala)
 FT /transl_except= (pos:195..197, aa:Thr)
 XX
 XX WO9749417-A1.
 PN
 XX 31-DEC-1997.
 PD
 XX 23-JUN-1997; 97WO-US10850.
 PF
 XX 24-JUN-1996; 96US-0669161.
 PR
 XX (LUDW-) LUDWIG INST CANCER RES.
 PA
 XX Boon-Falleur T, Debacker O, Van Den Eynde B;
 PI WPI; 1998-076905/07.
 DR P-PSDB; AAW47599.
 XX
 PT Isolated nucleic acid encoding GAGE tumour rejection antigen

ID	AAZ97216 standard; cDNA; 1024 BP.	Db	214 GAAGCAGCTTGAATGATGCAGGCTCTCTATGTTGGAATTTGTTCAATAAATTCT	155
XX	AAZ97216;	Qy	744 CTCAATAAAGTTTTCAGCTTTCACCAAAAAA 781	
AC		Db	154 CCAATAAAGCTTACAGCTTCTGTAATAAAAAA 117	
XX	18-APR-2000 (first entry)			
DE	Human prostate cancer differentially expressed gene #77.	RESULT 40		
XX	Prostate cancer specific gene; cancer; tumour progression; diagnosis;	AAID14982		
KW	hyperproliferative cell growth; prostatic disorder; treatment;	ID	AAID14982 standard; DNA; 1051 BP.	
KW	metastatic prostate cancer; benign prostate hyperplasia; BPH; ss.	XX		
XX		AC	AAID14982;	
OS	Homo sapiens.	XX		
XX	WO964594-A2.	DT	01-NOV-2001 (first entry)	
PN	16-DEC-1999.	XX		
PD		DE	Human NOV3 DNA.	
XX		XX		
PF	10-JUN-1999; 99WO-US13181.	KW	Human; NOVX; G-antigen; GAGE-like protein; interferon;	
XX		KW	G-protein coupled receptor; GPCR; hepatocyte nuclear factor;	
PR	11-JUN-1998; 98US-0088877.	KW	mast cell protease; gene therapy; proliferative disorder; cancer;	
PR	09-JUN-1999; 99US-0088877.	KW	immune disorder; hepatic disorder; cirrhosis; viral infection;	
XX	(CHIR) CHIRON CORP.	KW	hepatitis; neuroofactory system-related disorder; neurological disorder;	
PA		KW	Parkinson's disease; infertility; autoimmune disease; arthritis;	
XX		KW	multiple sclerosis; allergy; wound healing; cytostatic; nootropic;	
PI	Astel JH, Carroll E, Endege WO, Ford DM, Monahan JE, Schlegel R;	KW	immunosuppressive; neuroprotective; vulnerary; hepatotropic; ds.	
PI	Steinmann KE, Zhang J;	OS	Homo sapiens.	
XX	WPI; 2000-116541/10.	XX		
XX		FH	Key	Location/Qualifiers
XX	New isolated prostate cancer specific nucleic acids, used to develop	FT	5'UTR	1..592
PT	products for the diagnosis and treatment of cancer .	FT		/tag= a
XX		FT	CDS	593..928
XX		FT		/tag= b
PS	Claim 2; Page 107; 212pp; English.	FT	3'UTR	/product= "Human NOV3 protein"
XX		FT		929..1051
CC	This sequence represents a prostate cancer specific nucleic acid	FT		/tag= c
CC	sequence. The invention relates to a method for diagnosing cancer,	XX	WO200161009-A2.	
CC	tumour progression, hyperproliferative cell growth or accompanying	PN		
CC	biological and physical manifestations. The method involves contacting	XX	23-AUG-2001.	
CC	the biological sample with a probe that comprises a sequence capable of	PD		
CC	hybridising to any of the 339 nucleotide sequences given in the	XX	15-FEB-2001; 2001WO-US04828.	
CC	specification (see AA97140-297478) and detecting duplex formation. The	PF		
CC	products and methods of the invention can be used for the diagnosis,	XX	15-FEB-2000; 2000US-0182723.	
CC	prognosis, and treatment of cancer, tumour progression,	PR	15-FEB-2000; 2000US-0182724.	
CC	hyperproliferative cell growth, and accompanying physical and biological	PR	15-FEB-2000; 2000US-0182733.	
CC	manifestations. They can be used particularly for prostatic disorders	PR	22-FEB-2000; 2000US-0183896.	
CC	such as metastatic prostate cancer, localised prostate cancer, or benign	PR	23-FEB-2000; 2000US-0184275.	
CC	prostate hyperplasia (BPH).	PR	23-FEB-2000; 2000US-0184482.	
XX		PR	24-FEB-2000; 2000US-0184744.	
SQ	Sequence 1024 BP; 187 A; 290 C; 232 G; 270 T; 45 other;	PR	10-AUG-2000; 2000US-0197083.	
		PR	18-SEP-2000; 2000US-0224157.	
		PR	10-AUG-2000; 2000US-0231405.	
		PR	27-SEP-2000; 2000US-0236060.	
		PR	02-JAN-2001; 2001US-0259414.	
		PR	18-JAN-2001; 2001US-0262454.	
		PR	14-FEB-2001; 2001US-0783429.	
		XX	(CURA-) CURAGEN CORP.	
Qy	447 CCCCAAAAGAACCCAGCAGCTGAAAGTGGGATCTACACTGGGCGAGACAGAA 506	PA		
Db	450 CCTGAAGAGGGGAACCACTCAACGCTCAGGATCTCGAGCTGCTCAGGAG--GGA 394	XX		
Qy	507 GAAGTCAGGATACAGCTGAGATCCAGTCGCCACATGGAAGTGATCTGCAAGAGCTG 566	XX		
Db	393 GAGGATGAGGAGGATNTGCGAGTCAAGGGCCGAAGCTGAAGCTCATAGCCAGGAACAG 334	PI	Malyankar UM, Tchernev VT, Padigaru M, Taupier RJ, Spytek KA;	
Qy	567 CATCAGTCAAAACCCGGGATAAATCTGGATTTGGGTTCCGGGCTCAAGGTGAAGATAAT 626	PI	Majumder K, Guo X, Spaderna SK, Boldog FL;	
Db	333 GGTCAACCCACAGACTGGGTGTGAGTGTGAAGAT--GGTCTGTGGCGAGGATGGACCC 275	XX	WPI; 2001-514775/56.	
Qy	627 ACCTTAACAGCAACTGTAAATGCCAGAGCAGGTGAAGACCAACCAAGTTTAAAT 686	DR	P-PSDB; AAE08582.	
Db	274 GCCAAATCCAGAGAGGTGAAGACCGCTGAAGAGGTGAAGAACCAATCACAGTGTAAAA 215	XX		
Qy	687 GAAGCAAGCTGAACCAACCAAGCTGGTTTATATTAGATATTG---ACTTAAACTAT 743	PT	Isolated novel polypeptides useful for diagnosis of and treating	
		PT	cancer, infertility, autoimmune diseases, arthritis, multiple	
		PT	sclerosis, allergies, wound healing and hepatic disorders -	
		XX	Claim 9; Page 11; 140pp; English.	
		XX		

The present sequence is a human NOV3 DNA. The NOVX protein has homology with one of G-antigen (GAGE)-like protein, interferon, G-protein coupled receptor (GPCR), hepatocyte nuclear factor or mast cell protease. The NOVX is useful for treating or preventing a pathology associated with NOVX. It is also useful for determining the presence or amount of NOVX DNA in a sample, for identifying a potential therapeutic agent and in gene therapy. It is also useful for determining the presence of or predisposition to a disease associated with altered levels of NOVX. It is also useful for the diagnosis and treatment of proliferative disorders, e.g., cancer, immune disorders, hepatic disorders, e.g., cirrhosis, viral infections, e.g., hepatitis, neurofascory system-related disorders, neurological disorders, e.g., parkinson's disease, infertility, autoimmune diseases, arthritis, multiple sclerosis, allergies and wound healing.

Sequence 1051 BP: 317 A: 192 C: 257 G: 285 T: 0 other: XX

	Query Match	13.2%;	Score 103.2;	DB 22;	Length 1051;
	Best Local Similarity	62.8%;	Pred. No. 1.5e-20;		
	Matches 230;	Conservative	0;	Mismatches 123;	Indels 13; Gaps 4;
Qy	425	CCAGGAGCCCTAGTAAATGGAGAGCCCAAAAAGAGAAGAACACGACAGCTCAAACTCGGGATCC	484		
Db	679	CCAGGAGCCCTAGTAAATGGAGAGCCCAAAAAGAGAAGAACACGACAGCTCAAACTCGGGATCC	738		
Qy	485	TACACTGGGCGAGCAGACAGAGAAGATACAGGTACAGCTGAGATCCGACGTCCGCGACAT	544		
Db	739	TGCACCTAGTGGGAGATTGAAA--ATCAAGCAGTGCCTGCTTTTCAAGGGCCTGACAT	795		
Qy	545	GGAGGTGATCTGCAAGAGCTGCATCAGTCAACACCGGGGATAAATCTGGATTTGGGTT	604		
Db	796	GGAGGCTTTTCAACAGGAACCTGGCTCTGTTAAATAGAGGATGAGGCTGGAGAT-GGTC	854		
Qy	605	CCGGCGTCAAGGTGAAGATAATACCTAAAGAGGAACACTGTAAATGCCAAGCAGGTG	664		
Db	855	CTGATGTCAGGGGGTATTATGCCCACTTTTGATCTCACTAAAGTCTGNAAGCAGGTG	914		
Qy	665	AAGAGCAACCAACAGTTTAAATGAAGACAA-----GCTGAACACGCAAGCTGGTTTT	718		
Db	915	ATGCGCAACCATAGGTTTCAAGCAAGACAAATGAAGACTGAAGACCAAGAACGTTATTCTT	974		
Qy	719	ATATTAGATATTTCAGCTTTAAACTA---TCTCAATAAAAGTTTTTGCAGCTTTTCAACCAAAAA	775		
Db	975	AATCTGGAATTTTACTGATATAATATCTCTTAATAAAGTTTTTAAGTTTTCTGCAAGAAA	1034		
Qy	776	AAAAA	781		
Db	1035	AAAAA	1040		

Search completed: October 17, 2002, 00:29:01
Job time : 209 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 16, 2002, 21:53:26 ; Search time 47 Seconds
(without alignments)
4081.697 Million cell up

Title: US-09-658-824-808
Perfect score: 781
Sequence: 1 ggcgcgcgaactgtgaccqg.....ctttcaccaaaaaaaaaa 781

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 10

Total number of hits satisfying chosen parameters: 89210

```
Minimum DB seq length: 0
Maximum DB seq length: 20
```

Post-processing: Listing first 1000 summaries

```
Database : Issued_Patents_NA:**
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:**
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:**
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:**
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:**
5: /cgn2_6/ptodata/2/ina/PTCUS_COMB.seq:**
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:**
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %		DB	ID	Description
		Match	Length			
c 1	19	2.4	705	4	US-09-328-111-674	Sequence 674, Appl
c 2	19	2.4	1547	1	US-08-377-292-2	Sequence 1, Appli
c 3	19	2.4	1607	1	US-07-679-451-1	Sequence 1, Appli
c 4	19	2.4	1607	2	US-07-989-847-1	Sequence 3, Appli
c 5	19	2.4	1607	3	US-07-721-847A-3	Sequence 1, Appli
c 6	19	2.4	1607	4	US-08-469-411-1	Sequence 1, Appli
c 7	19	2.4	1607	4	US-08-925-779-3	Sequence 3, Appli
c 8	19	2.4	1607	6	516058-3	Patent No. 516058
c 9	19	2.4	2017	1	US-07-667-276A-3	Sequence 3, Appli
10	19	2.4	2328	4	US-08-811-481-34	Sequence 34, Appl
11	18	2.3	414	1	US-08-377-687-48	Sequence 48, Appl
12	18	2.3	414	1	US-08-777-192-48	Sequence 48, Appl
13	18	2.3	414	4	US-08-971-982-48	Sequence 48, Appl
14	18	2.3	3585	3	US-08-549-846-2	Sequence 2, Appli
15	18	2.3	5137	5	PCR-US96-01314-39	Sequence 39, Appl
16	18	2.3	5138	2	US-08-476-062A-39	Sequence 39, Appl
17	18	2.3	12537	2	US-08-611-280-4	Sequence 4, Appli
18	17	2.2	12537	4	US-09-195-940-4	Sequence 4, Appli
19	17	2.2	708	3	US-08-955-937A-3	Sequence 3, Appli
20	17	2.2	708	4	US-09-300-985-3	Sequence 3, Appli
21	17	2.2	1396	1	US-08-123-161A-11	Sequence 11, Appl
22	17	2.2	1396	1	US-08-483-278-11	Sequence 11, Appl
23	17	2.2	1529	3	US-09-189-760-5	Sequence 5, Appli
24	17	2.2	1529	3	US-09-188-811-5	Sequence 5, Appli
25	17	2.2	1529	4	US-09-514-423-5	Sequence 5, Appli
c 26	17	2.2	1770	2	US-08-749-289-2	Sequence 2, Appli
27	17	2.2	2080	2	US-08-878-563A-2	Sequence 2, Appli

101	16	2.0	2403	2	US-08-469-334-30	Sequence 30, Appl	c 174	15	1.9	575	1	US-08-507-016-8	Sequence 8, Appl
102	16	2.0	2403	3	US-09-300-529-30	Sequence 30, Appl	c 175	15	1.9	598	4	US-09-247-155-42	Sequence 42, Appl
103	16	2.0	2403	3	US-09-233-336A-7	Sequence 7, Appl	c 176	15	1.9	611	4	US-09-385-982-393	Sequence 393, App
104	16	2.0	2403	3	US-09-233-752A-7	Sequence 7, Appl	c 177	15	1.9	617	4	US-09-385-982-294	Sequence 294, App
105	16	2.0	2403	4	US-09-402-036-7	Sequence 7, Appl	c 178	15	1.9	622	4	US-09-385-982-57	Sequence 57, Appl
106	16	2.0	2502	3	US-09-234-332-1	Sequence 1, Appl	c 179	15	1.9	627	4	US-09-328-111-93	Sequence 93, Appl
107	16	2.0	3225	1	US-08-306-691B-45	Sequence 45, Appl	c 180	15	1.9	635	1	US-08-455-633A-35	Sequence 35, Appl
108	16	2.0	3225	5	PCT-US93-06251-91	Sequence 91, Appl	c 181	15	1.9	635	2	US-08-416-336-5	Sequence 35, Appl
109	16	2.0	3230	5	PCT-US93-06251-89	Sequence 89, Appl	c 182	15	1.9	635	5	PCT-US94-05354-35	Sequence 327, App
110	16	2.0	3306	1	US-08-261-206A-71	Sequence 71, Appl	c 183	15	1.9	635	4	US-08-385-982-327	Sequence 49, Appl
111	16	2.0	3728	1	US-08-111-939-15	Sequence 1, Appl	c 184	15	1.9	658	4	US-08-896-164-49	Sequence 2, Appl
112	16	2.0	5408	1	US-08-441-139-15	Sequence 15, Appl	c 185	15	1.9	731	1	US-08-451-405A-2	Sequence 173, App
113	16	2.0	5775	1	US-08-306-691B-15	Sequence 15, Appl	c 186	15	1.9	778	3	US-08-961-083-173	Sequence 47, Appl
114	16	2.0	5775	5	PCT-US93-06251-29	Sequence 29, Appl	c 187	15	1.9	790	1	US-08-306-691B-47	Sequence 93, Appl
115	16	2.0	6463	2	US-08-962-284-3	Sequence 3, Appl	c 188	15	1.9	790	5	PCT-US93-06251-93	Sequence 1, Appl
116	16	2.0	11298	1	US-07-869-933-31	Sequence 31, Appl	c 189	15	1.9	792	3	US-08-950-720A-1	Sequence 9, Appl
117	16	2.0	11298	1	US-08-201-879A-2	Sequence 2, Appl	c 190	15	1.9	807	2	US-08-531-927B-9	Sequence 1, Appl
118	16	2.0	11298	4	US-09-103-663-31	Sequence 31, Appl	c 191	15	1.9	812	3	US-08-462-778-1	Sequence 10, Appl
119	16	2.0	17949	4	US-09-087-465-3	Sequence 3, Appl	c 192	15	1.9	826	1	US-08-631-200-10	Sequence 10, Appl
120	16	2.0	31571	1	US-08-323-443B-1	Sequence 1, Appl	c 193	15	1.9	826	1	US-08-829-553-10	Sequence 10, Appl
121	16	2.0	53526	3	US-08-658-136-2	Sequence 2, Appl	c 194	15	1.9	826	2	US-08-922-267A-10	Sequence 10, Appl
122	16	2.0	53577	3	US-08-658-136-1	Sequence 1, Appl	c 195	15	1.9	826	2	US-08-936-707A-10	Sequence 10, Appl
123	16	2.0	80246	4	US-09-078-294-4	Sequence 4, Appl	c 196	15	1.9	826	3	US-08-936-706A-10	Sequence 10, Appl
124	16	2.0	80595	4	US-09-078-294-3	Sequence 3, Appl	c 197	15	1.9	826	3	US-09-248-203-10	Sequence 10, Appl
125	16	2.0	4403765	4	US-09-103-840A-2	Sequence 2, Appl	c 198	15	1.9	826	4	US-09-406-071-10	Sequence 10, Appl
126	16	2.0	4411529	4	US-09-103-840A-1	Sequence 1, Appl	c 199	15	1.9	854	4	US-09-064-693A-24	Sequence 24, Appl
127	15	1.9	15	3	US-08-832-021-52	Sequence 52, Appl	c 200	15	1.9	856	1	US-08-374-983A-12	Sequence 12, Appl
128	15	1.9	24	2	US-08-704-682-1	Sequence 1, Appl	c 201	15	1.9	857	4	US-08-998-416-555	Sequence 555, App
129	15	1.9	30	3	US-08-836-329-9	Sequence 9, Appl	c 202	15	1.9	866	4	US-09-227-357-97	Sequence 97, Appl
130	15	1.9	38	1	US-08-222-177A-68	Sequence 68, Appl	c 203	15	1.9	911	4	US-09-364-230-25	Sequence 25, Appl
131	15	1.9	40	4	US-09-306-290-2	Sequence 2, Appl	c 204	15	1.9	912	3	US-08-688-988-3	Sequence 3, Appl
132	15	1.9	40	4	US-09-306-290-5	Sequence 5, Appl	c 205	15	1.9	948	1	US-09-247-155-50	Sequence 50, Appl
133	15	1.9	40	4	US-09-306-290-9	Sequence 9, Appl	c 206	15	1.9	960	4	US-07-639-330E-1	Sequence 1, Appl
134	15	1.9	40	4	US-08-150-156A-37	Sequence 37, Appl	c 207	15	1.9	1020	4	US-09-247-155-143	Sequence 143, App
135	15	1.9	42	4	US-08-306-290-13	Sequence 13, Appl	c 208	15	1.9	1029	4	US-09-077-675A-4	Sequence 4, Appl
136	15	1.9	44	1	US-08-222-177A-249	Sequence 249, App	c 209	15	1.9	1048	4	US-09-227-357-127	Sequence 127, App
137	15	1.9	51	1	US-08-222-177A-325	Sequence 325, App	c 210	15	1.9	1048	4	US-07-639-330E-2	Sequence 2, Appl
138	15	1.9	51	2	US-08-704-682-2	Sequence 2, Appl	c 211	15	1.9	1095	1	US-08-389-668A-3	Sequence 3, Appl
139	15	1.9	72	1	US-08-222-177A-131	Sequence 131, App	c 212	15	1.9	1106	1	US-08-732-506-3	Sequence 3, Appl
140	15	1.9	72	1	US-08-222-177A-427	Sequence 427, App	c 213	15	1.9	1106	5	PCT-US95-05768-3	Sequence 5, Appl
141	15	1.9	75	2	US-08-776-944-13	Sequence 13, App	c 214	15	1.9	1109	4	US-08-664-962B-5	Sequence 5, Appl
142	15	1.9	92	1	US-08-222-177A-430	Sequence 430, App	c 215	15	1.9	1109	4	US-09-311-743-5	Sequence 5, Appl
143	15	1.9	98	1	US-08-399-412A-77	Sequence 77, Appl	c 216	15	1.9	1115	1	US-08-784-651-9	Sequence 9, Appl
144	15	1.9	130	6	5198345-15	Patent No. 5198345	c 217	15	1.9	1117	2	US-08-960-022-1	Sequence 1, Appl
145	15	1.9	194	1	US-08-222-177A-15	Sequence 15, Appl	c 218	15	1.9	1188	2	US-08-557-128-10	Sequence 10, Appl
146	15	1.9	216	4	US-09-439-313-425	Sequence 425, App	c 219	15	1.9	1191	1	US-08-687-379-13	Sequence 13, Appl
147	15	1.9	228	3	US-08-688-988-25	Sequence 25, App	c 220	15	1.9	1191	1	US-08-687-379-15	Sequence 15, Appl
148	15	1.9	294	1	US-08-446-660-18	Sequence 18, App	c 221	15	1.9	1212	3	US-09-058-489-43	Sequence 43, Appl
149	15	1.9	294	4	US-08-974-302-18	Sequence 18, App	c 222	15	1.9	1220	1	US-08-374-983A-14	Sequence 14, Appl
150	15	1.9	300	1	US-08-333-358-5	Sequence 5, Appl	c 223	15	1.9	1234	1	US-08-798-000-1	Sequence 1, Appl
151	15	1.9	300	1	US-08-463-694-5	Sequence 5, Appl	c 224	15	1.9	1260	1	US-08-599-252-79	Sequence 79, Appl
152	15	1.9	300	1	US-08-694-501-5	Sequence 5, Appl	c 225	15	1.9	1260	1	US-08-436-074-52	Sequence 52, Appl
153	15	1.9	301	5	PCT-US94-05150-20	Sequence 20, App	c 226	15	1.9	1260	5	PCT-US96-06352-79	Sequence 79, Appl
154	15	1.9	332	2	US-08-469-412A-8	Sequence 8, Appl	c 227	15	1.9	1260	5	PCT-US96-06583-79	Sequence 79, Appl
155	15	1.9	332	3	US-09-021-715-8	Sequence 8, Appl	c 228	15	1.9	1268	4	US-09-303-064-23	Sequence 23, Appl
156	15	1.9	342	3	US-09-014-877A-7	Sequence 7, Appl	c 229	15	1.9	1268	4	US-09-086-503-23	Sequence 23, Appl
157	15	1.9	390	4	US-09-385-982-232	Sequence 232, App	c 230	15	1.9	1289	4	US-09-247-155-138	Sequence 138, App
158	15	1.9	446	1	US-08-104-072B-1	Sequence 1, Appl	c 231	15	1.9	1315	4	US-08-992-035A-2	Sequence 2, Appl
159	15	1.9	474	2	US-08-619-542B-45	Sequence 45, App	c 232	15	1.9	1333	4	US-09-247-155-77	Sequence 77, Appl
160	15	1.9	492	2	US-08-892-880-12	Sequence 12, App	c 233	15	1.9	1372	6	5189147-2	Patent No. 5189147
161	15	1.9	513	4	US-09-268-364-7	Sequence 7, Appl	c 234	15	1.9	1374	1	US-08-278-630A-9	Sequence 9, Appl
162	15	1.9	533	1	US-08-104-073-1	Sequence 1, Appl	c 235	15	1.9	1380	2	US-08-467-559B-1	Sequence 1, Appl
163	15	1.9	534	1	US-08-599-252-101	Sequence 101, App	c 236	15	1.9	1387	3	US-08-868-594-1	Sequence 1, Appl
164	15	1.9	534	5	PCT-US96-06352-101	Sequence 101, App	c 237	15	1.9	1387	3	US-09-434-323-1	Sequence 4, Appl
165	15	1.9	534	5	PCT-US96-06583-101	Sequence 101, App	c 238	15	1.9	1430	2	US-08-204-288-4	Sequence 4, Appl
166	15	1.9	535	4	US-09-385-982-385	Sequence 385, App	c 239	15	1.9	1512	3	US-08-909-965C-8	Sequence 8, Appl
167	15	1.9	536	4	US-09-385-982-347	Sequence 347, App	c 240	15	1.9	1519	3	US-08-592-900-1	Sequence 1, Appl
168	15	1.9	558	4	US-08-976-259-53	Sequence 53, App	c 241	15	1.9	1566	4	US-09-227-357-141	Sequence 141, App
169	15	1.9	560	2	US-08-365-486A-7	Sequence 7, Appl	c 242	15	1.9	1567	4	US-09-276-531-75	Sequence 75, App
170	15	1.9	560	4	US-08-880-342-7	Sequence 7, Appl	c 243	15	1.9	1597	2	US-08-724-974A-1	Sequence 1, Appl
171	15	1.9	566	4	US-09-385-982-468	Sequence 468, App	c 244	15	1.9	1619	3	US-09-163-162-1	Sequence 1, Appl
172	15	1.9	569	4	US-09-227-357-89	Sequence 89, App	c 245	15	1.9	1619	4	US-09-286-407-1	Sequence 1, Appl
173	15	1.9	574	4	US-09-385-982-463	Sequence 463, App	c 246	15	1.9	1619	4	US-09-286-407-1	Sequence 1, Appl

c 247	15	1.9	1619	4	US-09-496-694B-97	Sequence 97, Appl	320	15	1.9	2584	5	PCT-US93-08322-1	Sequence 1, Appli
248	15	1.9	1632	2	US-08-892-715-1	Sequence 1, Appli	321	15	1.9	2610	2	US-08-989-386-2	Sequence 2, Appli
249	15	1.9	1632	2	US-09-145-947-1	Sequence 1, Appli	322	15	1.9	2628	1	US-08-143-219-1	Sequence 1, Appli
250	15	1.9	1632	4	US-09-265-642-1	Sequence 1, Appli	323	15	1.9	2647	5	PCT-US93-06251-77	Sequence 77, Appl
251	15	1.9	1633	1	US-08-197-792-42	Sequence 42, Appl	324	15	1.9	2657	2	US-08-592-363-3	Sequence 3, Appli
252	15	1.9	1633	1	US-08-459-850-42	Sequence 42, Appl	325	15	1.9	2658	2	US-08-826-422-14	Sequence 14, Appl
253	15	1.9	1633	1	US-08-459-214-42	Sequence 42, Appl	326	15	1.9	2659	3	US-08-749-522-1	Sequence 1, Appli
254	15	1.9	1644	2	US-08-458-555-1	Sequence 1, Appli	327	15	1.9	2684	2	US-08-984-171-2	Sequence 2, Appli
255	15	1.9	1648	4	US-09-303-064-26	Sequence 26, Appl	328	15	1.9	2688	2	US-08-909-965C-1	Sequence 1, Appli
256	15	1.9	1648	4	US-09-086-503-26	Sequence 26, Appl	329	15	1.9	2738	2	US-08-795-868-17	Sequence 17, Appl
257	15	1.9	1679	4	US-09-271-437-3	Sequence 3, Appli	330	15	1.9	2738	4	US-09-303-069-17	Sequence 17, Appl
258	15	1.9	1708	3	US-09-108-020-5	Sequence 5, Appli	331	15	1.9	2770	4	US-08-426-509A-5	Sequence 5, Appli
259	15	1.9	1715	1	US-07-847-743B-24	Sequence 24, Appl	332	15	1.9	2770	5	PCT-US95-05008-5	Sequence 31, Appl
260	15	1.9	1715	1	US-08-456-201-24	Sequence 24, Appl	333	15	1.9	2811	1	US-08-040-548-31	Sequence 31, Appl
261	15	1.9	1715	2	US-08-456-241-24	Sequence 24, Appl	334	15	1.9	2811	1	US-08-466-344-31	Sequence 31, Appl
262	15	1.9	1715	5	PCT-US92-04235A-24	Sequence 24, Appl	335	15	1.9	2817	6	5206152-6	Patent No. 5206152
c 263	15	1.9	1724	4	US-09-153-804-8	Sequence 8, Appli	336	15	1.9	2928	2	US-08-095-728B-3	Sequence 3, Appli
264	15	1.9	1733	3	US-09-147-522-1	Sequence 1, Appli	337	15	1.9	2928	5	PCT-US92-02320A-3	Sequence 3, Appli
265	15	1.9	1776	4	US-08-531-927B-1	Sequence 1, Appli	338	15	1.9	2940	2	US-08-592-383-1	Sequence 1, Appli
266	15	1.9	1791	1	US-09-041-886-12	Sequence 12, Appl	339	15	1.9	2940	6	5171671-1	Patent No. 5171671
267	15	1.9	1791	1	US-08-245-294-7	Sequence 7, Appli	340	15	1.9	2946	3	US-08-968-563-6	Sequence 6, Appli
268	15	1.9	1791	1	US-08-474-499-7	Sequence 7, Appli	341	15	1.9	2946	3	US-08-969-683A-6	Sequence 2, Appli
269	15	1.9	1791	1	US-08-307-279A-7	Sequence 7, Appli	342	15	1.9	2946	4	US-09-297-928-2	Sequence 2, Appli
270	15	1.9	1791	5	PCT-US95-06211-7	Sequence 7, Appli	343	15	1.9	2977	2	US-08-820-170A-36	Sequence 36, Appl
271	15	1.9	1840	2	US-08-484-993B-15	Sequence 15, Appl	344	15	1.9	2977	3	US-09-055-699-36	Sequence 36, Appl
272	15	1.9	1840	2	US-08-484-158B-15	Sequence 15, Appl	345	15	1.9	2977	4	US-09-273-565-36	Sequence 36, Appl
273	15	1.9	1840	2	US-08-484-596A-15	Sequence 15, Appl	346	15	1.9	2977	4	US-09-565-538-36	Sequence 36, Appl
274	15	1.9	1840	2	US-08-480-150A-15	Sequence 15, Appl	347	15	1.9	3046	1	US-08-726-725-1	Sequence 1, Appli
275	15	1.9	1840	3	US-08-458-731-15	Sequence 15, Appl	348	15	1.9	3113	1	US-08-146-422-20	Sequence 20, Appl
276	15	1.9	1840	3	US-08-149-223A-15	Sequence 15, Appl	349	15	1.9	3113	1	US-08-626-554-2	Sequence 2, Appli
277	15	1.9	1847	3	US-08-930-894-3	Sequence 3, Appli	350	15	1.9	3132	3	US-09-286-904-45	Sequence 45, Appl
278	15	1.9	1854	3	US-08-923-454A-13	Sequence 13, Appl	351	15	1.9	3134	2	US-08-533-689A-1	Sequence 1, Appli
c 279	15	1.9	1857	2	US-08-537-715-3	Sequence 3, Appli	352	15	1.9	3219	2	US-08-687-289A-4	Sequence 3, Appli
280	15	1.9	1857	5	PCT-US94-04173-3	Sequence 3, Appli	353	15	1.9	3219	2	US-08-968-751-3	Sequence 3, Appli
c 281	15	1.9	1857	5	US-09-377-557-9	Sequence 9, Appli	354	15	1.9	3256	2	US-08-968-751-3	Sequence 3, Appli
282	15	1.9	1863	3	US-09-199-229-1	Sequence 1, Appli	355	15	1.9	3358	3	US-08-957-063-4	Sequence 4, Appli
283	15	1.9	1863	3	US-09-443-087-1	Sequence 1, Appli	356	15	1.9	3358	3	US-09-487-685-4	Sequence 4, Appli
284	15	1.9	1863	4	US-09-687-298-1	Sequence 1, Appli	357	15	1.9	3366	1	US-08-469-802B-1	Sequence 1, Appli
c 285	15	1.9	1864	1	US-08-454-720A-38	Sequence 38, Appl	358	15	1.9	3366	2	US-08-267-803B-1	Sequence 29, Appl
286	15	1.9	1894	2	US-08-935-450-7	Sequence 7, Appli	359	15	1.9	3376	1	US-08-320-559-29	Sequence 29, Appl
287	15	1.9	1965	1	US-08-258-420-9	Sequence 9, Appli	360	15	1.9	3376	5	PCT-US94-04496-29	Sequence 29, Appl
288	15	1.9	1975	1	US-08-484-105-11	Sequence 11, Appl	361	15	1.9	3376	5	PCT-US94-04496-29	Sequence 29, Appl
289	15	1.9	1975	1	US-08-484-106-11	Sequence 11, Appl	362	15	1.9	3384	2	US-08-687-289A-1	Sequence 1, Appli
290	15	1.9	2001	3	US-08-850-961-13	Sequence 13, Appl	363	15	1.9	3387	1	US-08-468-557-1	Sequence 1, Appli
291	15	1.9	2001	4	US-09-479-776-13	Sequence 13, Appl	364	15	1.9	3393	1	US-08-295-502-1	Sequence 1, Appli
292	15	1.9	2124	4	US-09-198-122-1	Sequence 1, Appli	365	15	1.9	3393	5	PCT-US95-10691-1	Sequence 6, Appli
293	15	1.9	2152	1	US-08-188-582-17	Sequence 17, Appl	366	15	1.9	3396	2	US-08-974-565C-6	Sequence 6, Appli
294	15	1.9	2162	2	US-08-646-715-17	Sequence 5, Appli	367	15	1.9	3396	3	US-09-255-748-6	Sequence 1, Appli
c 295	15	1.9	2186	3	US-08-959-382-1	Sequence 1, Appli	368	15	1.9	3402	1	US-08-480-917-1	Sequence 1, Appli
296	15	1.9	2201	2	US-08-974-565C-2	Sequence 2, Appli	369	15	1.9	3402	4	US-09-138-736-1	Sequence 1, Appli
297	15	1.9	2201	3	US-09-255-748-2	Sequence 2, Appli	370	15	1.9	3402	4	US-09-138-736-1	Sequence 1, Appli
298	15	1.9	2214	4	US-08-943-731-57	Sequence 57, Appl	371	15	1.9	3402	1	US-08-253-155A-7	Sequence 7, Appli
299	15	1.9	2226	2	US-08-031-538-1	Sequence 1, Appli	372	15	1.9	3407	1	US-08-389-564B-1	Sequence 1, Appli
300	15	1.9	2226	2	US-08-624-663A-1	Sequence 1, Appli	373	15	1.9	3461	2	US-08-466-047B-1	Sequence 1, Appli
301	15	1.9	2229	3	US-08-892-880-1	Sequence 1, Appli	374	15	1.9	3471	3	US-09-527-236A-1	Sequence 1, Appli
302	15	1.9	2313	2	US-08-393-333-1	Sequence 1, Appli	375	15	1.9	3474	4	US-09-303-069-24	Sequence 24, Appl
303	15	1.9	2385	1	US-08-891-845-11	Sequence 11, Appl	376	15	1.9	3487	4	US-09-303-069-24	Sequence 3, Appli
304	15	1.9	2387	3	US-09-011-074-3	Sequence 3, Appli	377	15	1.9	3490	2	US-08-841-483-3	Sequence 3, Appli
c 305	15	1.9	2418	3	US-08-820-170A-35	Sequence 35, Appl	378	15	1.9	3490	4	US-09-382-911-3	Sequence 3, Appli
306	15	1.9	2430	3	US-09-055-699-35	Sequence 35, Appl	379	15	1.9	3600	3	US-08-894-731-1	Sequence 1, Appli
c 307	15	1.9	2430	3	US-09-273-565-35	Sequence 35, Appl	380	15	1.9	3714	4	US-09-087-465-9	Sequence 9, Appli
308	15	1.9	2430	4	US-09-565-538-35	Sequence 35, Appl	381	15	1.9	3804	2	US-08-483-488-5	Sequence 5, Appli
c 309	15	1.9	2467	1	US-08-148-5810-1	Sequence 1, Appli	382	15	1.9	4094	2	US-08-841-483-5	Sequence 5, Appli
310	15	1.9	2467	1	US-07-943-843-3	Sequence 3, Appli	383	15	1.9	4094	4	US-08-952-089A-29	Sequence 29, Appl
311	15	1.9	2498	1	US-08-347-003-3	Sequence 3, Appli	384	15	1.9	4146	4	US-08-041-538-1	Sequence 1, Appli
312	15	1.9	2498	1	US-08-931-608A-2	Sequence 2, Appli	385	15	1.9	4300	1	US-08-463-642-1	Sequence 1, Appli
313	15	1.9	2520	4	US-08-448-603A-31	Sequence 31, Appl	386	15	1.9	4300	1	US-08-465-602-1	Sequence 1, Appli
314	15	1.9	2570	3	US-08-121-714-1	Sequence 1, Appli	387	15	1.9	4300	2	PCT-US91-09422-1	Sequence 1, Appli
315	15	1.9	2570	3	US-08-322-742-13	Sequence 13, Appl	388	15	1.9	4300	5	US-08-470-350B-1	Sequence 3, Appli
316	15	1.9	2584	1	US-08-477-108A-1	Sequence 1, Appli	389	15	1.9	4360	1	US-07-885-972A-3	Sequence 3, Appli
317	15	1.9	2584	1	US-08-477-108A-1	Sequence 1, Appli	390	15	1.9	4405	2	US-08-745-880-3	Sequence 3, Appli
318	15	1.9	2584	2	US-08-477-112-1	Sequence 1, Appli	391	15	1.9	4405	2	US-08-745-880-3	Sequence 3, Appli
319	15	1.9	2584	2	US-08-477-112-1	Sequence 1, Appli	392	15	1.9	4405	2	US-08-745-880-3	Sequence 3, Appli

393	15	1.9	4405	2	US-08-480-382-3	Sequence 3, Appli	c 466	15	1.9	17410	2	US-08-449-700-3	Sequence 3, Appli
394	15	1.9	4437	1	US-08-559-303B-72	Sequence 72, Appl	c 467	15	1.9	17410	2	US-08-449-699A-3	Sequence 3, Appli
395	15	1.9	4437	1	US-09-175-828-72	Sequence 72, Appl	c 468	15	1.9	17415	3	US-08-486-343A-1	Sequence 1, Appli
396	15	1.9	4543	2	US-08-519-547A-5	Sequence 5, Appli	c 469	15	1.9	17415	5	PCT-US95-07349-1	Sequence 1, Appli
397	15	1.9	4698	1	US-07-807-043B-5	Sequence 5, Appli	c 470	15	1.9	18609	4	US-08-943-731-1	Sequence 1, Appli
398	15	1.9	4698	1	US-08-299-849B-5	Sequence 5, Appli	c 471	15	1.9	28720	4	US-09-341-587-7	Sequence 7, Appli
399	15	1.9	4698	2	US-08-142-368A-5	Sequence 5, Appli	c 472	15	1.9	29604	3	US-08-781-891-207	Sequence 207, App
400	15	1.9	4698	3	US-08-967-727-5	Sequence 5, Appli	c 473	15	1.9	35060	3	US-08-814-095-7	Sequence 7, Appli
401	15	1.9	4698	4	US-08-037-230B-5	Sequence 5, Appli	c 474	15	1.9	35100	5	US-08-306-691B-19	Sequence 19, Appl
402	15	1.9	4718	3	US-08-936-135-8	Sequence 8, Appli	c 475	15	1.9	35100	5	PCT-US93-06251-19	Sequence 101, App
403	15	1.9	4724	3	US-09-194-285-8	Sequence 11, Appl	c 476	15	1.9	43795	4	US-08-742-185-101	Sequence 17, Appl
404	15	1.9	4733	3	US-08-936-135-11	Sequence 11, Appl	c 477	15	1.9	48974	4	US-08-920-422-17	Sequence 17, Appl
405	15	1.9	4749	1	US-08-452-259-1	Sequence 1, Appli	c 478	15	1.9	48974	4	US-08-920-422-17	Sequence 3, Appli
406	15	1.9	4749	5	PCT-US96-07336-1	Sequence 13, Appl	c 479	15	1.9	59065	4	US-09-813-817-3	Sequence 1, Appli
407	15	1.9	4769	3	US-08-936-135-13	Sequence 15, Appl	c 480	15	1.9	72928	3	US-09-009-913-1	Sequence 1, Appli
408	15	1.9	4784	3	US-08-936-135-15	Sequence 15, Appl	c 481	15	1.9	80161	3	US-09-036-987A-1	Sequence 1, Appli
409	15	1.9	4821	3	US-08-913-374-1	Sequence 4, Appli	c 482	15	1.9	80161	4	US-09-370-700-1	Sequence 3, Appli
410	15	1.9	5110	2	US-08-404-531B-4	Sequence 4, Appli	c 483	15	1.9	80246	4	US-09-078-294-4	Sequence 3, Appli
411	15	1.9	5110	2	US-08-404-531B-5	Sequence 4, Appli	c 484	15	1.9	80246	4	US-09-078-294-4	Sequence 3, Appli
412	15	1.9	5110	3	US-08-476-900A-5	Sequence 5, Appli	c 485	15	1.9	84495	4	US-09-797-906-3	Sequence 16, Appl
413	15	1.9	5110	3	US-08-488-546A-4	Sequence 4, Appli	c 486	15	1.9	84495	4	US-09-797-906-3	Sequence 16, Appl
414	15	1.9	5110	3	US-08-488-546A-5	Sequence 5, Appli	c 487	15	1.9	152331	3	US-09-128-155-16	Sequence 17, Appl
415	15	1.9	5110	3	US-08-488-546A-5	Sequence 5, Appli	c 488	15	1.9	152331	3	US-09-128-155-16	Sequence 17, Appl
416	15	1.9	5236	5	PCT-US91-09422-16	Sequence 16, Appl	c 489	15	1.9	246240	2	US-08-724-394A-20	Sequence 20, Appl
417	15	1.9	5261	1	US-08-045-806-3	Sequence 3, Appli	c 490	15	1.9	246240	2	US-08-724-394A-21	Sequence 21, Appl
418	15	1.9	5261	1	US-08-366-051B-3	Sequence 3, Appli	c 491	15	1.9	246240	2	US-08-724-394A-22	Sequence 22, Appl
419	15	1.9	5526	3	US-08-751-359-21	Sequence 21, Appl	c 492	15	1.9	246240	2	US-08-724-394A-22	Sequence 22, Appl
420	15	1.9	5526	3	US-08-907-146-21	Sequence 21, Appl	c 493	15	1.9	246240	2	US-08-724-394A-22	Sequence 22, Appl
421	15	1.9	5555	1	US-08-484-438-3	Sequence 3, Appli	c 494	15	1.9	246240	2	US-08-724-394A-22	Sequence 22, Appl
422	15	1.9	5993	4	US-09-383-630-1	Sequence 2, Appli	c 495	14	1.8	14	3	US-08-832-021-15	Sequence 15, Appl
423	15	1.9	5993	4	US-09-383-630-2	Sequence 2, Appli	c 496	14	1.8	14	3	US-08-832-021-15	Sequence 15, Appl
424	15	1.9	6028	4	US-09-011-745-5	Sequence 8, Appli	c 497	14	1.8	14	3	US-08-832-021-15	Sequence 15, Appl
425	15	1.9	6350	2	US-08-385-735A-8	Sequence 8, Appli	c 498	14	1.8	16	1	US-08-284-484A-4	Sequence 12, Appl
426	15	1.9	6370	4	US-09-245-041-12	Sequence 12, Appl	c 499	14	1.8	20	4	US-09-228-942-7	Sequence 12, Appl
427	15	1.9	6639	1	US-08-727-034-1	Sequence 2, Appli	c 500	14	1.8	20	4	US-08-482-918-32	Sequence 32, Appl
428	15	1.9	6641	4	US-09-064-693A-25	Sequence 25, Appl	c 501	14	1.8	20	4	US-09-224-681-32	Sequence 32, Appl
429	15	1.9	6642	1	US-08-727-034-5	Sequence 5, Appli	c 502	14	1.8	20	4	US-08-336-728A-32	Sequence 4, Appli
430	15	1.9	6669	3	US-09-212-971-5	Sequence 5, Appli	c 503	14	1.8	21	1	US-08-128-011-4	Sequence 142, App
431	15	1.9	6669	3	US-08-800-929A-5	Sequence 5, Appli	c 504	14	1.8	21	2	US-08-403-888A-142	Sequence 146, App
432	15	1.9	6669	4	US-09-617-053A-5	Sequence 5, Appli	c 505	14	1.8	21	2	US-08-403-888A-146	Sequence 23, Appl
433	15	1.9	6843	1	US-08-727-034-6	Sequence 6, Appli	c 506	14	1.8	22	1	US-08-235-180-23	Sequence 5, Appli
434	15	1.9	6911	1	US-08-311-174-4	Sequence 4, Appli	c 507	14	1.8	22	1	US-08-332-420-5	Sequence 1, Appli
435	15	1.9	6961	1	US-08-727-034-2	Sequence 2, Appli	c 508	14	1.8	22	3	US-08-974-180-1	Sequence 1, Appli
436	15	1.9	7070	1	US-08-619-554-3	Sequence 3, Appli	c 509	14	1.8	23	5	PCT-US94-05407-7	Sequence 8, Appli
437	15	1.9	7452	3	US-08-592-500-1	Sequence 1, Appli	c 510	14	1.8	23	5	PCT-US94-05407-8	Sequence 8, Appli
438	15	1.9	7452	3	US-08-195-006-1	Sequence 1, Appli	c 511	14	1.8	24	1	US-09-434-131A-6	Sequence 5, Appli
439	15	1.9	7452	5	PCT-US94-07644A-1	Sequence 1, Appli	c 512	14	1.8	25	1	US-08-683-877-5	Sequence 624, App
440	15	1.9	8083	4	US-09-383-630-4	Sequence 5, Appli	c 513	14	1.8	26	3	US-08-388-353-624	Sequence 624, App
441	15	1.9	8083	4	US-09-383-630-5	Sequence 6, Appli	c 514	14	1.8	26	3	US-08-488-551B-624	Sequence 5169, Ap
442	15	1.9	8342	3	US-08-545-860D-63	Sequence 63, Appl	c 515	14	1.8	27	4	US-08-584-040-5169	Sequence 410, App
443	15	1.9	8342	5	PCT-US94-04496-63	Sequence 63, Appl	c 516	14	1.8	29	2	US-08-559-998-410	Sequence 410, App
444	15	1.9	8387	2	US-08-532-814-1	Sequence 1, Appli	c 517	14	1.8	29	4	US-09-225-928-410	Sequence 114, App
445	15	1.9	8388	4	US-09-225-509-1	Sequence 14, Appl	c 518	14	1.8	30	4	US-08-018-584A-114	Sequence 8, Appli
446	15	1.9	8589	4	US-09-245-041-14	Sequence 1, Appli	c 519	14	1.8	33	1	US-08-255-670A-8	Sequence 18, Appl
447	15	1.9	8827	4	US-09-245-041-1	Sequence 3, Appli	c 520	14	1.8	37	4	US-08-133-711-18	Sequence 4, Appli
448	15	1.9	8878	1	US-08-206-176-3	Sequence 3, Appli	c 521	14	1.8	40	4	US-09-434-131A-4	Sequence 4, Appli
449	15	1.9	9620	4	US-08-952-127-11	Sequence 11, Appl	c 522	14	1.8	40	2	US-08-306-290-4	Sequence 11, Appl
450	15	1.9	9934	4	US-08-977-171-2	Sequence 2, Appli	c 523	14	1.8	42	3	US-08-477-459-11	Sequence 11, Appl
451	15	1.9	10718	3	US-08-325-426B-1	Sequence 21, Appl	c 524	14	1.8	42	3	US-08-479-869-11	Sequence 11, Appl
452	15	1.9	10815	4	US-09-004-838-21	Sequence 1, Appli	c 525	14	1.8	42	3	US-08-486-414-11	Sequence 11, Appl
453	15	1.9	11288	4	US-08-646-301A-1	Sequence 1, Appli	c 526	14	1.8	42	5	PCT-US94-01826A-11	Sequence 11, Appl
454	15	1.9	11288	4	US-08-481-968A-4	Sequence 4, Appli	c 527	14	1.8	42	5	PCT-US94-02252A-11	Sequence 40, Appli
455	15	1.9	11288	4	US-08-154-712B-4	Sequence 4, Appli	c 528	14	1.8	46	1	US-08-232-463-40	Sequence 9, Appli
456	15	1.9	11461	3	US-08-669-161A-29	Sequence 29, Appl	c 529	14	1.8	46	5	PCT-US91-02954-9	Sequence 3, Appli
457	15	1.9	13149	4	US-09-004-838-87	Sequence 87, Appl	c 530	14	1.8	46	5	PCT-US94-05407-3	Sequence 10, Appl
458	15	1.9	14796	4	US-08-975-080-35	Sequence 35, Appl	c 531	14	1.8	46	5	PCT-US94-05407-10	Sequence 273, App
459	15	1.9	14796	4	US-09-630-706-10	Sequence 10, Appl	c 532	14	1.8	47	4	US-09-338-907-273	Sequence 273, App
460	15	1.9	14796	4	US-09-496-694B-3	Sequence 3, Appli	c 533	14	1.8	47	4	US-09-218-207-273	Sequence 70, Appl
461	15	1.9	15144	3	US-08-458-434A-6	Sequence 6, Appli	c 534	14	1.8	49	3	US-08-651-472-70	Sequence 70, Appl
462	15	1.9	17056	4	US-09-245-041-3	Sequence 3, Appli	c 535	14	1.8	49	4	US-08-358-928-70	Sequence 39, Appl
463	15	1.9	17410	1	US-07-841-646-3	Sequence 3, Appli	c 536	14	1.8	49	5	PCT-US96-07709-39	Sequence 40, Appl
464	15	1.9	17410	1	US-08-147-023-3	Sequence 3, Appli	c 537	14	1.8	49	5	PCT-US96-07709-40	Sequence 40, Appl
465	15	1.9	17410	1	US-08-447-570-3	Sequence 3, Appli	c 538	14	1.8	49	5	PCT-US96-07709-40	Sequence 40, Appl

c 539	14	1.8	49	5	PCT-US96-09848-19	Sequence 19, Appl	c 612	14	1.8	108	4	US-08-375-992A-32	Sequence 32, Appl
c 540	14	1.8	49	5	PCT-US96-09848-20	Sequence 56, Appl	c 613	14	1.8	108	5	PCT-US93-00324-32	Sequence 32, Appl
c 541	14	1.8	51	3	US-08-651-472-56	Sequence 56, Appl	c 614	14	1.8	110	2	US-08-602-093-19	Sequence 19, Appl
c 542	14	1.8	51	3	US-08-358-928-56	Sequence 56, Appl	c 615	14	1.8	111	2	US-08-097-554A-78	Sequence 78, Appl
c 543	14	1.8	53	3	US-08-651-472-57	Sequence 57, Appl	c 616	14	1.8	111	2	US-08-097-554A-78	Sequence 78, Appl
c 544	14	1.8	53	4	US-08-358-928-57	Sequence 57, Appl	c 617	14	1.8	111	2	US-08-912-129A-6	Sequence 6, Appl
c 545	14	1.8	54	4	US-08-584-040-8265	Sequence 8265, Ap	c 618	14	1.8	114	3	US-09-035-648-13	Sequence 13, Appl
c 546	14	1.8	55	1	US-07-803-633A-3	Sequence 3, Appl	c 619	14	1.8	114	3	US-09-001-951-13	Sequence 13, Appl
c 547	14	1.8	55	1	US-07-803-633A-4	Sequence 4, Appl	c 620	14	1.8	117	3	US-08-480-640A-78	Sequence 78, Appl
c 548	14	1.8	55	1	US-07-750-080A-38	Sequence 38, Appl	c 621	14	1.8	117	3	US-08-295-802-78	Sequence 78, Appl
c 549	14	1.8	55	2	US-07-750-080A-39	Sequence 39, Appl	c 622	14	1.8	117	4	US-08-686-968C-130	Sequence 130, App
c 550	14	1.8	55	2	US-08-525-742-13	Sequence 14, Appl	c 623	14	1.8	117	4	US-08-686-968C-144	Sequence 144, App
c 551	14	1.8	55	2	US-08-525-742-14	Sequence 14, Appl	c 624	14	1.8	117	4	US-08-488-237A-78	Sequence 78, Appl
c 552	14	1.8	55	3	US-08-651-472-38	Sequence 38, Appl	c 625	14	1.8	119	2	US-08-375-992A-86	Sequence 86, Appl
c 553	14	1.8	55	3	US-08-651-472-39	Sequence 38, Appl	c 626	14	1.8	119	2	US-08-097-554A-86	Sequence 86, Appl
c 554	14	1.8	55	4	US-08-358-928-38	Sequence 38, Appl	c 627	14	1.8	119	2	US-08-097-554A-95	Sequence 95, Appl
c 555	14	1.8	55	4	US-08-358-928-39	Sequence 39, Appl	c 628	14	1.8	119	2	US-08-097-554A-95	Sequence 95, Appl
c 556	14	1.8	56	2	US-08-776-944-9	Sequence 9, Appl	c 629	14	1.8	119	2	US-08-659-251-49	Sequence 49, Appl
c 557	14	1.8	59	4	US-08-952-793-236	Sequence 236, App	c 630	14	1.8	119	4	US-09-256-490-49	Sequence 49, Appl
c 558	14	1.8	59	5	PCT-US96-09455A-236	Sequence 236, App	c 631	14	1.8	120	5	PCT-US96-11445-49	Sequence 49, Appl
c 559	14	1.8	60	4	US-09-396-281-8	Sequence 6, Appl	c 632	14	1.8	120	4	US-08-686-968C-151	Sequence 151, App
c 560	14	1.8	61	2	US-09-032-684-6	Sequence 8, Appl	c 633	14	1.8	120	4	US-08-686-968C-140	Sequence 140, App
c 561	14	1.8	63	1	US-08-232-463-47	Sequence 8, Appl	c 634	14	1.8	123	3	US-08-480-640A-86	Sequence 86, Appl
c 562	14	1.8	69	4	US-09-269-911A-8	Sequence 43, Appl	c 635	14	1.8	123	3	US-08-480-640A-86	Sequence 86, Appl
c 563	14	1.8	70	2	US-08-097-554A-43	Sequence 43, Appl	c 636	14	1.8	124	4	US-08-295-802-86	Sequence 95, Appl
c 564	14	1.8	70	3	US-08-480-640A-43	Sequence 43, Appl	c 637	14	1.8	124	4	US-08-488-237A-86	Sequence 86, Appl
c 565	14	1.8	70	3	US-08-295-802-43	Sequence 43, Appl	c 638	14	1.8	124	4	US-08-375-992A-86	Sequence 86, Appl
c 566	14	1.8	70	4	US-08-488-237A-43	Sequence 43, Appl	c 639	14	1.8	124	4	US-08-375-992A-95	Sequence 95, Appl
c 567	14	1.8	70	4	US-08-375-992A-43	Sequence 43, Appl	c 640	14	1.8	127	1	US-07-750-080A-13	Sequence 13, Appl
c 568	14	1.8	74	2	US-08-097-554A-44	Sequence 44, Appl	c 641	14	1.8	127	3	US-08-480-640A-160	Sequence 160, App
c 569	14	1.8	74	3	US-08-480-640A-44	Sequence 44, Appl	c 642	14	1.8	127	3	US-08-651-472-13	Sequence 13, Appl
c 570	14	1.8	74	3	US-08-295-802-44	Sequence 44, Appl	c 643	14	1.8	127	3	US-08-295-802-160	Sequence 160, App
c 571	14	1.8	74	4	US-08-488-237A-44	Sequence 44, Appl	c 644	14	1.8	127	4	US-08-488-237A-160	Sequence 160, App
c 572	14	1.8	74	4	US-08-375-992A-44	Sequence 44, Appl	c 645	14	1.8	127	4	US-08-358-928-13	Sequence 13, Appl
c 573	14	1.8	80	1	US-07-920-281C-25	Sequence 25, Appl	c 646	14	1.8	127	4	US-08-375-992A-160	Sequence 160, App
c 574	14	1.8	80	4	US-08-466-277-25	Sequence 25, Appl	c 647	14	1.8	128	3	US-08-480-640A-145	Sequence 145, App
c 575	14	1.8	86	2	US-08-525-742-50	Sequence 50, Appl	c 648	14	1.8	128	3	US-08-295-802-145	Sequence 145, App
c 576	14	1.8	88	1	US-07-750-080A-11	Sequence 11, Appl	c 649	14	1.8	128	4	US-08-488-237A-145	Sequence 145, App
c 577	14	1.8	88	3	US-08-651-472-11	Sequence 11, Appl	c 650	14	1.8	128	4	US-08-375-992A-145	Sequence 145, App
c 578	14	1.8	88	4	US-08-358-928-11	Sequence 11, Appl	c 651	14	1.8	133	3	US-08-480-640A-176	Sequence 176, App
c 579	14	1.8	91	2	US-08-525-742-51	Sequence 51, Appl	c 652	14	1.8	133	3	US-08-295-802-176	Sequence 176, App
c 580	14	1.8	92	1	US-07-750-080A-12	Sequence 12, Appl	c 653	14	1.8	133	4	US-08-488-237A-176	Sequence 176, App
c 581	14	1.8	92	3	US-08-651-472-12	Sequence 12, Appl	c 654	14	1.8	133	4	US-08-375-992A-176	Sequence 176, App
c 582	14	1.8	92	4	US-08-358-928-12	Sequence 12, Appl	c 655	14	1.8	134	1	US-07-750-080A-14	Sequence 14, Appl
c 583	14	1.8	93	1	US-07-750-080A-40	Sequence 40, Appl	c 656	14	1.8	134	3	US-08-651-472-14	Sequence 14, Appl
c 584	14	1.8	93	3	US-08-525-742-28	Sequence 28, Appl	c 657	14	1.8	134	4	US-08-358-928-14	Sequence 14, Appl
c 585	14	1.8	93	3	US-08-651-472-40	Sequence 40, Appl	c 658	14	1.8	137	2	US-08-486-414-37	Sequence 37, Appl
c 586	14	1.8	93	4	US-08-358-928-40	Sequence 40, Appl	c 659	14	1.8	137	3	US-08-486-414-37	Sequence 37, Appl
c 587	14	1.8	95	2	US-08-525-742-29	Sequence 29, Appl	c 660	14	1.8	138	3	US-08-480-640A-198	Sequence 198, App
c 588	14	1.8	96	2	US-08-525-742-30	Sequence 30, Appl	c 661	14	1.8	138	4	US-08-686-968C-203	Sequence 203, App
c 589	14	1.8	97	1	US-07-750-080A-41	Sequence 41, Appl	c 662	14	1.8	138	4	US-08-488-237A-198	Sequence 198, App
c 590	14	1.8	97	2	US-08-525-742-27	Sequence 27, Appl	c 663	14	1.8	138	4	US-08-375-992A-198	Sequence 198, App
c 591	14	1.8	97	3	US-08-651-472-41	Sequence 41, Appl	c 664	14	1.8	140	3	US-08-480-640A-181	Sequence 181, App
c 592	14	1.8	97	4	US-08-358-928-41	Sequence 41, Appl	c 665	14	1.8	140	3	US-08-295-802-181	Sequence 181, App
c 593	14	1.8	97	4	US-08-952-793-205	Sequence 205, App	c 666	14	1.8	140	4	US-08-488-237A-181	Sequence 181, App
c 594	14	1.8	97	5	PCT-US96-09455A-205	Sequence 205, App	c 667	14	1.8	140	4	US-08-375-992A-181	Sequence 181, App
c 595	14	1.8	98	4	US-08-991-789A-282	Sequence 282, App	c 668	14	1.8	141	3	US-08-480-640A-203	Sequence 203, App
c 596	14	1.8	98	4	US-09-062-451-282	Sequence 282, App	c 669	14	1.8	141	3	US-08-686-968C-198	Sequence 198, App
c 597	14	1.8	102	1	US-07-820-154A-10	Sequence 10, Appl	c 670	14	1.8	141	4	US-08-488-237A-203	Sequence 203, App
c 598	14	1.8	102	2	US-08-097-554A-10	Sequence 10, Appl	c 671	14	1.8	141	4	US-08-375-992A-203	Sequence 203, App
c 599	14	1.8	102	3	US-08-480-640A-10	Sequence 10, Appl	c 672	14	1.8	141	4	US-08-488-237A-203	Sequence 203, App
c 600	14	1.8	102	3	US-08-295-802-10	Sequence 10, Appl	c 673	14	1.8	150	4	US-09-439-313-320	Sequence 320, App
c 601	14	1.8	102	4	US-08-686-968C-106	Sequence 106, App	c 674	14	1.8	153	2	US-08-484-575A-27	Sequence 27, Appl
c 602	14	1.8	102	4	US-08-991-789A-272	Sequence 272, App	c 675	14	1.8	153	3	US-08-486-414-42	Sequence 42, Appl
c 603	14	1.8	102	4	US-08-488-237A-10	Sequence 10, Appl	c 676	14	1.8	156	4	US-09-019-095A-41	Sequence 41, Appl
c 604	14	1.8	102	4	US-08-375-992A-10	Sequence 10, Appl	c 677	14	1.8	162	3	US-08-746-411A-3	Sequence 3, Appl
c 605	14	1.8	102	4	US-09-062-451-272	Sequence 10, Appl	c 678	14	1.8	168	3	US-08-480-640A-150	Sequence 150, App
c 606	14	1.8	102	5	PCT-US93-00324-10	Sequence 10, Appl	c 679	14	1.8	168	3	US-08-295-802-150	Sequence 150, App
c 607	14	1.8	108	1	US-07-820-154A-32	Sequence 32, Appl	c 680	14	1.8	168	4	US-08-686-968C-120	Sequence 120, App
c 608	14	1.8	108	2	US-08-097-554A-32	Sequence 32, Appl	c 681	14	1.8	168	4	US-08-488-237A-120	Sequence 120, App
c 609	14	1.8	108	3	US-08-480-640A-32	Sequence 32, Appl	c 682	14	1.8	168	4	US-08-375-992A-150	Sequence 150, App
c 610	14	1.8	108	3	US-08-295-802-32	Sequence 32, Appl	c 683	14	1.8	177	4	US-08-686-968C-147	Sequence 147, App
c 611	14	1.8	108	4	US-08-488-237A-32	Sequence 32, Appl	c 684	14	1.8	178	2	US-08-484-575A-16	Sequence 16, Appl

c 685	14	1.8	178	3	US-08-477-459-16	Sequence 16, Appl	c 758	14	1.8	300	2	US-09-032-684-17	Sequence 17, Appl
c 686	14	1.8	178	3	US-08-479-869-16	Sequence 16, Appl	c 759	14	1.8	302	1	US-08-189-199A-1	Sequence 1, Appl
c 687	14	1.8	178	3	US-08-486-414-16	Sequence 16, Appl	c 760	14	1.8	303	1	US-08-358-171-23	Sequence 23, Appl
c 688	14	1.8	178	5	PCT-US94-01826A-16	Sequence 16, Appl	c 761	14	1.8	304	3	US-09-090-947-23	Sequence 23, Appl
c 689	14	1.8	178	5	PCT-US94-02252A-16	Sequence 16, Appl	c 762	14	1.8	308	3	US-09-035-648-16	Sequence 16, Appl
c 690	14	1.8	180	2	US-08-097-554A-72	Sequence 72, Appl	c 763	14	1.8	308	4	US-09-001-951-16	Sequence 16, Appl
c 691	14	1.8	180	3	US-08-480-640A-72	Sequence 72, Appl	c 764	14	1.8	313	3	US-08-398-633-8	Sequence 8, Appl
c 692	14	1.8	180	3	US-08-295-802-72	Sequence 72, Appl	c 765	14	1.8	314	4	US-08-976-259-41	Sequence 41, Appl
c 693	14	1.8	180	4	US-08-686-968C-116	Sequence 116, Appl	c 766	14	1.8	315	4	US-08-945-424-1	Sequence 1, Appl
c 694	14	1.8	180	4	US-08-488-237A-72	Sequence 72, Appl	c 767	14	1.8	322	4	US-09-385-982-216	Sequence 216, App
c 695	14	1.8	180	4	US-08-375-992A-72	Sequence 72, Appl	c 768	14	1.8	322	4	US-09-385-982-362	Sequence 362, App
c 696	14	1.8	182	2	US-08-097-554A-64	Sequence 64, Appl	c 769	14	1.8	326	3	US-09-035-648-9	Sequence 9, Appl
c 697	14	1.8	182	2	US-08-484-575A-15	Sequence 15, Appl	c 770	14	1.8	326	3	US-09-001-951-9	Sequence 9, Appl
c 698	14	1.8	182	3	US-08-477-459-15	Sequence 15, Appl	c 771	14	1.8	327	3	US-08-766-355-9	Sequence 9, Appl
c 699	14	1.8	182	3	US-08-480-640A-64	Sequence 64, Appl	c 772	14	1.8	327	3	US-09-385-982-544	Sequence 544, App
c 700	14	1.8	182	3	US-08-479-869-15	Sequence 15, Appl	c 773	14	1.8	327	4	US-09-003-198A-9	Sequence 9, Appl
c 701	14	1.8	182	3	US-08-295-802-64	Sequence 64, Appl	c 774	14	1.8	327	4	US-09-428-805-9	Sequence 9, Appl
c 702	14	1.8	182	3	US-08-486-414-15	Sequence 15, Appl	c 775	14	1.8	333	4	US-09-018-584A-27	Sequence 27, Appl
c 703	14	1.8	182	4	US-08-991-789A-289	Sequence 289, App	c 776	14	1.8	336	2	US-08-691-814B-96	Sequence 96, Appl
c 704	14	1.8	182	4	US-08-488-237A-64	Sequence 64, Appl	c 777	14	1.8	342	4	US-09-385-982-342	Sequence 342, App
c 705	14	1.8	182	4	US-08-375-992A-64	Sequence 64, Appl	c 778	14	1.8	348	1	US-08-686-878A-54	Sequence 54, Appl
c 706	14	1.8	182	4	US-09-062-451-289	Sequence 289, App	c 779	14	1.8	348	2	US-08-721-924-4	Sequence 4, Appl
c 707	14	1.8	182	5	PCT-US94-01826A-15	Sequence 15, Appl	c 780	14	1.8	348	4	US-09-385-982-283	Sequence 283, App
c 708	14	1.8	182	5	PCT-US94-02252A-15	Sequence 15, Appl	c 781	14	1.8	348	4	US-09-175-928-8	Sequence 8, Appl
c 709	14	1.8	185	3	US-08-480-640A-156	Sequence 156, App	c 782	14	1.8	349	3	US-09-087-194-37	Sequence 37, Appl
c 710	14	1.8	185	3	US-08-295-802-156	Sequence 156, App	c 783	14	1.8	352	4	US-09-385-982-529	Sequence 529, App
c 711	14	1.8	185	4	US-08-488-237A-156	Sequence 156, App	c 784	14	1.8	374	2	US-08-370-156-24	Sequence 24, Appl
c 712	14	1.8	185	4	US-08-375-992A-156	Sequence 156, App	c 785	14	1.8	374	4	US-09-385-982-135	Sequence 135, App
c 713	14	1.8	186	4	US-08-686-968C-132	Sequence 132, App	c 786	14	1.8	380	4	US-09-385-982-457	Sequence 457, App
c 714	14	1.8	186	4	US-08-686-968C-137	Sequence 137, App	c 787	14	1.8	382	3	US-09-035-648-10	Sequence 10, Appl
c 715	14	1.8	186	4	US-08-686-968C-154	Sequence 154, App	c 788	14	1.8	382	4	US-09-001-951-10	Sequence 10, Appl
c 716	14	1.8	187	3	US-09-094-287-7	Sequence 7, Appl	c 789	14	1.8	390	4	US-08-905-223-98	Sequence 98, Appl
c 717	14	1.8	189	4	US-09-113-750A-29	Sequence 29, Appl	c 790	14	1.8	400	2	US-08-244-537-3	Sequence 3, Appl
c 718	14	1.8	191	2	US-09-032-684-4	Sequence 4, Appl	c 791	14	1.8	403	3	US-09-035-648-3	Sequence 3, Appl
c 719	14	1.8	193	3	US-08-480-640A-171	Sequence 171, App	c 792	14	1.8	403	4	US-09-385-982-29	Sequence 29, Appl
c 720	14	1.8	193	3	US-08-295-802-171	Sequence 171, App	c 793	14	1.8	403	4	US-09-001-951-3	Sequence 3, Appl
c 721	14	1.8	193	4	US-08-752-540-6	Sequence 6, Appl	c 794	14	1.8	404	4	US-09-386-493-5	Sequence 5, Appl
c 722	14	1.8	193	4	US-08-488-237A-171	Sequence 171, App	c 795	14	1.8	405	1	US-08-469-427A-8	Sequence 8, Appl
c 723	14	1.8	193	4	US-08-375-992A-171	Sequence 171, App	c 796	14	1.8	405	2	US-08-609-443B-8	Sequence 8, Appl
c 724	14	1.8	198	3	US-09-035-648-5	Sequence 5, Appl	c 797	14	1.8	405	2	US-08-569-063C-8	Sequence 8, Appl
c 725	14	1.8	198	4	US-08-686-968C-125	Sequence 125, App	c 798	14	1.8	405	4	US-09-020-956-101	Sequence 101, App
c 726	14	1.8	198	4	US-09-001-951-5	Sequence 5, Appl	c 799	14	1.8	405	4	US-09-030-607-101	Sequence 101, App
c 727	14	1.8	201	4	US-09-113-750A-28	Sequence 28, Appl	c 800	14	1.8	405	4	US-09-439-313-101	Sequence 101, App
c 728	14	1.8	206	1	US-07-820-154A-21	Sequence 21, Appl	c 801	14	1.8	411	4	US-09-385-982-5	Sequence 5, Appl
c 729	14	1.8	206	2	US-08-097-554A-21	Sequence 21, Appl	c 802	14	1.8	419	4	US-09-030-607-201	Sequence 201, App
c 730	14	1.8	206	3	US-08-480-640A-21	Sequence 21, Appl	c 803	14	1.8	419	4	US-09-439-313-201	Sequence 201, App
c 731	14	1.8	206	3	US-08-295-802-21	Sequence 21, Appl	c 804	14	1.8	421	3	US-08-983-409-2	Sequence 2, Appl
c 732	14	1.8	206	4	US-08-686-968C-111	Sequence 111, App	c 805	14	1.8	425	2	US-08-967-101-102	Sequence 102, App
c 733	14	1.8	206	4	US-08-488-237A-21	Sequence 21, Appl	c 806	14	1.8	425	2	US-08-592-541-102	Sequence 102, App
c 734	14	1.8	206	4	US-08-375-992A-21	Sequence 21, Appl	c 807	14	1.8	425	3	US-09-124-698-102	Sequence 102, App
c 735	14	1.8	206	5	PCT-US93-00324-21	Sequence 21, Appl	c 808	14	1.8	425	4	US-09-127-480-102	Sequence 102, App
c 736	14	1.8	213	3	US-08-513-974B-37	Sequence 37, Appl	c 809	14	1.8	425	4	US-08-496-841C-102	Sequence 102, App
c 737	14	1.8	218	3	US-09-035-648-14	Sequence 14, Appl	c 810	14	1.8	426	2	US-08-975-316-52	Sequence 52, Appl
c 738	14	1.8	218	4	US-09-001-951-14	Sequence 14, Appl	c 811	14	1.8	430	4	US-08-991-789A-250	Sequence 250, App
c 739	14	1.8	219	4	US-08-991-789A-286	Sequence 286, App	c 812	14	1.8	430	4	US-09-062-451-250	Sequence 250, App
c 740	14	1.8	231	4	US-09-062-451-286	Sequence 286, App	c 813	14	1.8	431	4	US-08-991-789A-281	Sequence 281, App
c 741	14	1.8	231	4	US-09-439-313-463	Sequence 463, App	c 814	14	1.8	431	4	US-09-062-451-281	Sequence 281, App
c 742	14	1.8	234	1	US-08-840-683-3	Sequence 3, Appl	c 815	14	1.8	436	1	US-08-181-271A-41	Sequence 41, Appl
c 743	14	1.8	234	2	US-08-555-722-3	Sequence 3, Appl	c 816	14	1.8	436	1	US-08-449-315-41	Sequence 41, Appl
c 744	14	1.8	234	3	US-09-384-301-3	Sequence 318, App	c 817	14	1.8	436	1	US-08-444-803-41	Sequence 41, Appl
c 745	14	1.8	240	3	US-08-513-974B-318	Sequence 318, App	c 818	14	1.8	436	1	US-08-449-043-41	Sequence 41, Appl
c 746	14	1.8	249	3	US-08-480-640A-209	Sequence 209, App	c 819	14	1.8	436	1	US-08-456-265A-41	Sequence 41, Appl
c 747	14	1.8	249	3	US-08-686-968C-209	Sequence 209, App	c 820	14	1.8	436	1	US-08-455-416-41	Sequence 41, Appl
c 748	14	1.8	249	4	US-08-488-237A-209	Sequence 209, App	c 821	14	1.8	436	1	US-08-454-244-41	Sequence 41, Appl
c 749	14	1.8	249	4	US-08-375-992A-209	Sequence 209, App	c 822	14	1.8	436	1	US-08-454-876-41	Sequence 41, Appl
c 750	14	1.8	255	1	US-08-487-748A-2	Sequence 2, Appl	c 823	14	1.8	436	2	US-08-457-364-41	Sequence 41, Appl
c 751	14	1.8	255	3	US-08-398-633-2	Sequence 2, Appl	c 824	14	1.8	436	2	US-08-456-262-41	Sequence 41, Appl
c 752	14	1.8	255	3	US-08-480-070C-2	Sequence 2, Appl	c 825	14	1.8	436	2	US-08-456-240-41	Sequence 41, Appl
c 753	14	1.8	255	3	US-08-829-525-2	Sequence 2, Appl	c 826	14	1.8	436	2	US-08-455-736-41	Sequence 41, Appl
c 754	14	1.8	255	4	US-08-609-583A-2	Sequence 2, Appl	c 827	14	1.8	436	2	US-08-971-217-41	Sequence 41, Appl
c 755	14	1.8	255	4	US-08-937-399-2	Sequence 2, Appl	c 828	14	1.8	436	3	US-08-821-451A-3	Sequence 3, Appl
c 756	14	1.8	256	4	US-09-385-982-187	Sequence 187, App	c 829	14	1.8	436	4	US-09-263-810-3	Sequence 3, Appl
c 757	14	1.8	296	4	US-09-385-982-19	Sequence 19, Appl	c 830	14	1.8	436	4	US-09-350-600-41	Sequence 41, Appl

831	14	1.8	436	4	US-09-583-169-3	Sequence 3, Appl	904	14	1.8	584	4	US-09-347-801-11	Sequence 11, Appl
832	14	1.8	437	4	US-09-091-725-46	Sequence 46, Appl	905	14	1.8	585	2	US-08-882-704A-1	Sequence 1, Appl
c 833	14	1.8	440	2	US-08-967-101-105	Sequence 105, App	906	14	1.8	586	1	US-08-463-115-75	Sequence 75, Appl
c 834	14	1.8	440	2	US-08-967-101-105	Sequence 105, App	907	14	1.8	586	1	US-08-465-388-75	Sequence 75, Appl
c 835	14	1.8	440	3	US-09-552-341-105	Sequence 105, App	c 908	14	1.8	588	3	US-08-965-504B-1	Sequence 1, Appl
c 836	14	1.8	440	4	US-09-124-698-105	Sequence 105, App	909	14	1.8	591	1	US-08-090-523-24	Sequence 24, Appl
c 837	14	1.8	440	4	US-08-496-841C-105	Sequence 105, App	910	14	1.8	591	1	US-08-334-639-3	Sequence 3, Appl
c 838	14	1.8	450	2	US-08-967-101-7	Sequence 7, Appl	911	14	1.8	591	1	US-08-398-627-24	Sequence 24, Appl
c 839	14	1.8	450	2	US-08-592-541-7	Sequence 7, Appl	912	14	1.8	591	1	US-08-406-858-25	Sequence 25, Appl
c 840	14	1.8	450	3	US-09-124-698-7	Sequence 7, Appl	c 913	14	1.8	591	5	US-09-385-982-406	Sequence 406, pp
c 841	14	1.8	450	4	US-09-127-480-7	Sequence 7, Appl	914	14	1.8	591	5	PCT-US94-05275-25	Sequence 25, Appl
c 842	14	1.8	450	4	US-08-496-841C-7	Sequence 7, Appl	915	14	1.8	593	1	US-08-443-568B-13	Sequence 13, Appl
c 843	14	1.8	455	4	US-08-991-789A-273	Sequence 273, App	c 916	14	1.8	593	4	US-09-385-982-448	Sequence 448, App
c 844	14	1.8	455	4	US-09-062-451-273	Sequence 273, App	c 917	14	1.8	593	5	PCT-US94-06997-13	Sequence 13, Appl
c 845	14	1.8	461	4	US-08-991-789A-274	Sequence 274, App	918	14	1.8	602	4	US-09-040-984-11	Sequence 11, Appl
c 846	14	1.8	461	4	US-09-062-451-274	Sequence 274, App	919	14	1.8	602	4	US-09-123-912-11	Sequence 11, Appl
c 847	14	1.8	468	3	US-08-600-982-29	Sequence 29, Appl	920	14	1.8	603	2	US-08-924-838-6	Sequence 6, Appl
c 848	14	1.8	468	5	PCT-US94-10261A-29	Sequence 29, Appl	c 921	14	1.8	604	4	US-09-385-982-496	Sequence 496, App
c 849	14	1.8	470	4	US-09-020-956-102	Sequence 102, App	922	14	1.8	606	1	US-08-592-126-99	Sequence 99, Appl
c 850	14	1.8	470	4	US-09-030-607-102	Sequence 102, App	923	14	1.8	606	4	US-09-040-984-55	Sequence 55, Appl
c 851	14	1.8	470	4	US-09-439-313-102	Sequence 102, App	c 924	14	1.8	606	4	US-09-064-414-1	Sequence 1, Appl
c 852	14	1.8	479	4	US-09-328-111-432	Sequence 432, App	c 925	14	1.8	606	4	US-09-064-414-3	Sequence 3, Appl
c 853	14	1.8	485	1	US-08-324-243-37	Sequence 37, Appl	926	14	1.8	606	4	US-09-123-912-55	Sequence 55, Appl
c 854	14	1.8	485	1	US-08-532-390-37	Sequence 37, Appl	c 927	14	1.8	608	4	US-09-385-982-183	Sequence 183, App
c 855	14	1.8	485	3	US-08-717-324-37	Sequence 37, Appl	c 928	14	1.8	608	4	US-09-385-982-523	Sequence 523, App
c 856	14	1.8	485	5	PCT-US95-11511-37	Sequence 37, Appl	c 929	14	1.8	611	4	US-09-385-982-178	Sequence 178, App
c 857	14	1.8	490	4	US-09-003-198A-19	Sequence 19, Appl	c 930	14	1.8	613	4	US-09-385-982-144	Sequence 144, App
c 858	14	1.8	491	1	US-08-133-711-40	Sequence 40, Appl	931	14	1.8	614	4	US-09-385-982-255	Sequence 255, App
c 859	14	1.8	496	4	US-09-328-111-125	Sequence 125, App	932	14	1.8	615	4	US-09-385-982-240	Sequence 240, App
c 860	14	1.8	500	2	US-09-018-595B-3	Sequence 3, Appl	933	14	1.8	618	4	US-09-385-982-218	Sequence 218, App
c 861	14	1.8	500	3	US-09-324-709A-3	Sequence 3, Appl	c 934	14	1.8	619	4	US-09-328-111-829	Sequence 829, App
c 862	14	1.8	501	3	US-08-699-628-1	Sequence 1, Appl	c 935	14	1.8	622	4	US-09-385-982-184	Sequence 184, App
c 863	14	1.8	501	4	US-09-339-913B-96	Sequence 96, Appl	c 936	14	1.8	622	4	US-09-385-982-312	Sequence 312, App
c 864	14	1.8	501	4	US-09-339-904A-96	Sequence 96, Appl	937	14	1.8	623	4	US-08-905-223-25	Sequence 23, Appl
c 865	14	1.8	501	4	US-08-769-062B-96	Sequence 96, Appl	c 938	14	1.8	624	4	US-09-385-982-359	Sequence 359, App
c 866	14	1.8	501	4	US-09-344-002B-96	Sequence 96, Appl	c 939	14	1.8	626	4	US-09-385-982-39	Sequence 39, Appl
c 867	14	1.8	506	1	US-08-398-617-13	Sequence 13, Appl	940	14	1.8	626	4	US-09-064-414-5	Sequence 5, Appl
c 868	14	1.8	506	2	US-08-398-615-13	Sequence 13, Appl	c 941	14	1.8	627	4	US-09-385-982-39	Sequence 39, Appl
c 869	14	1.8	506	3	US-08-397-303-13	Sequence 13, Appl	c 942	14	1.8	628	4	US-09-385-982-505	Sequence 505, App
c 870	14	1.8	507	3	US-08-766-355-10	Sequence 10, Appl	c 943	14	1.8	629	4	US-09-385-982-305	Sequence 305, App
c 871	14	1.8	507	4	US-09-003-198A-10	Sequence 10, Appl	c 944	14	1.8	631	4	US-09-328-111-126	Sequence 126, App
c 872	14	1.8	507	4	US-09-428-805-10	Sequence 10, Appl	c 945	14	1.8	631	4	US-09-385-982-354	Sequence 354, App
c 873	14	1.8	509	4	US-09-030-607-202	Sequence 202, App	c 946	14	1.8	632	4	US-09-328-111-798	Sequence 798, App
c 874	14	1.8	509	4	US-09-439-313-202	Sequence 202, App	c 947	14	1.8	632	4	US-09-385-982-499	Sequence 499, App
c 875	14	1.8	515	3	US-08-545-809A-63	Sequence 63, Appl	948	14	1.8	634	2	US-08-222-719-15	Sequence 15, Appl
c 876	14	1.8	515	4	US-09-439-313-472	Sequence 472, App	949	14	1.8	634	2	US-08-470-925-15	Sequence 15, Appl
c 877	14	1.8	516	1	US-08-288-888-1	Sequence 1, Appl	950	14	1.8	634	2	US-08-471-613-15	Sequence 15, Appl
c 878	14	1.8	517	2	US-08-910-075-1	Sequence 1, Appl	951	14	1.8	634	5	PCT-US93-10443-15	Sequence 3, Appl
c 879	14	1.8	517	2	US-08-905-801A-1	Sequence 1, Appl	952	14	1.8	636	2	US-08-284-941-3	Sequence 3, Appl
c 880	14	1.8	528	1	US-08-422-613-1	Sequence 1, Appl	953	14	1.8	636	2	US-08-447-642-3	Sequence 3, Appl
c 881	14	1.8	532	3	US-09-036-113-1	Sequence 1, Appl	954	14	1.8	636	4	US-09-236-503-3	Sequence 3, Appl
c 882	14	1.8	535	3	US-08-094-287-8	Sequence 8, Appl	955	14	1.8	636	5	PCT-US93-02147A-3	Sequence 3, Appl
c 883	14	1.8	540	2	US-08-448-561-3	Sequence 3, Appl	c 956	14	1.8	647	4	US-08-927-219-44	Sequence 44, Appl
c 884	14	1.8	550	4	US-08-998-416-148	Sequence 148, App	957	14	1.8	651	2	US-08-912-129A-57	Sequence 57, Appl
c 885	14	1.8	552	4	US-09-020-956-79	Sequence 79, Appl	958	14	1.8	654	4	US-09-040-984-31	Sequence 31, Appl
c 886	14	1.8	552	4	US-09-030-607-79	Sequence 79, Appl	959	14	1.8	654	4	US-09-123-912-31	Sequence 31, Appl
c 887	14	1.8	552	4	US-09-439-313-79	Sequence 79, Appl	c 960	14	1.8	654	5	PCT-US91-02954-10	Sequence 10, Appl
c 888	14	1.8	557	4	US-09-385-982-395	Sequence 395, App	c 961	14	1.8	657	4	US-09-385-982-91	Sequence 91, Appl
c 889	14	1.8	558	4	US-09-030-607-181	Sequence 181, App	c 962	14	1.8	658	4	US-09-328-111-816	Sequence 816, App
c 890	14	1.8	558	4	US-09-439-313-181	Sequence 181, App	963	14	1.8	660	1	US-08-555-678-41	Sequence 41, Appl
c 891	14	1.8	564	1	US-08-117-362-32	Sequence 32, App	964	14	1.8	663	4	US-08-623-428D-6	Sequence 6, Appl
c 892	14	1.8	564	1	US-08-486-924-32	Sequence 32, App	c 965	14	1.8	669	4	US-09-328-111-782	Sequence 782, App
c 893	14	1.8	566	4	US-09-328-111-466	Sequence 466, App	c 966	14	1.8	675	4	US-08-998-416-1096	Sequence 1096, App
c 894	14	1.8	570	1	US-07-822-966B-1	Sequence 1, Appl	c 967	14	1.8	683	4	US-09-328-111-168	Sequence 168, App
c 895	14	1.8	571	1	US-08-322-742-14	Sequence 14, Appl	c 968	14	1.8	685	4	US-09-227-357-66	Sequence 66, Appl
c 896	14	1.8	578	4	US-09-328-111-757	Sequence 757, App	969	14	1.8	685	4	US-09-227-357-100	Sequence 100, App
c 897	14	1.8	580	2	US-08-809-763-2	Sequence 2, Appl	c 970	14	1.8	688	6	5498694-3	Patent No. 5498694
c 898	14	1.8	580	3	US-08-956-253-2	Sequence 2, Appl	c 971	14	1.8	690	4	US-09-328-111-74	Sequence 74, Appl
c 899	14	1.8	581	1	US-07-620-426B-30	Sequence 30, Appl	972	14	1.8	704	4	US-09-122-400B-8	Sequence 8, Appl
c 900	14	1.8	581	1	US-07-662-007B-36	Sequence 36, Appl	973	14	1.8	706	4	US-09-191-136-14	Sequence 14, Appl
c 901	14	1.8	581	1	US-07-824-247-36	Sequence 36, Appl	c 974	14	1.8	710	4	US-08-998-416-603	Sequence 603, App
c 902	14	1.8	581	3	US-08-470-204A-36	Sequence 36, Appl	c 975	14	1.8	715	4	US-08-901-789A-264	Sequence 264, App
c 903	14	1.8	581	4	US-09-385-982-12	Sequence 12, Appl	c 976	14	1.8	715	4	US-09-062-451-264	Sequence 264, App

5	977	14	1.8	718	4	US-09-123-912-96	Sequence 95, Appl 1
6	978	14	1.8	719	4	US-09-227-357-74	Sequence 74, Appl 1
7	979	14	1.8	720	4	US-08-998-416-957	Sequence 957, App
8	980	14	1.8	722	4	US-09-091-590A-1	Sequence 1, Appl 1
9	981	14	1.8	729	4	US-08-991-789A-275	Sequence 275, App
10	982	14	1.8	729	4	US-09-062-451-275	Sequence 275, App
11	983	14	1.8	730	4	US-08-991-789A-263	Sequence 263, App
12	984	14	1.8	730	4	US-09-062-451-263	Sequence 263, App
13	985	14	1.8	732	4	US-09-328-111-697	Sequence 697, App
14	986	14	1.8	735	3	US-08-950-720A-5	Sequence 5, Appl 1
15	987	14	1.8	741	2	US-08-912-129A-47	Sequence 47, Appl 1
16	988	14	1.8	750	3	US-09-167-717-3	Sequence 3, Appl 1
17	989	14	1.8	756	1	US-08-681-812-3	Sequence 3, Appl 1
18	990	14	1.8	757	1	US-07-808-451-11	Sequence 11, Appl 1
19	991	14	1.8	757	1	US-07-989-845-27	Sequence 27, Appl 1
20	992	14	1.8	757	1	US-07-989-844-11	Sequence 11, Appl 1
21	993	14	1.8	757	1	US-08-161-044-11	Sequence 11, Appl 1
22	994	14	1.8	757	1	US-08-240-121-11	Sequence 11, Appl 1
23	995	14	1.8	757	1	US-08-451-241-11	Sequence 11, Appl 1
24	996	14	1.8	757	5	PCT-US92-10625-11	Sequence 11, Appl 1
25	997	14	1.8	757	5	PCT-US93-11297-11	Sequence 11, Appl 1
26	998	14	1.8	757	5	PCT-US93-11298-27	Sequence 27, Appl 1
27	999	14	1.8	765	4	US-08-998-416-215	Sequence 215, App
28	1000	14	1.8	768	1	US-08-236-918A-5	Sequence 5, Appl 1

ALIGNMENTS

```

RESULT 1
US-09-328-111-674/c
; Sequence 674, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 674
; LENGTH: 705
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(705)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-111-674

```

Query Match 2.4%; Score 19; DB 4; Length 705;
Best Local Similarity 100.0%; Pred. No. 5.1;

QY 763 TTTCACCAAAAAAAAAAAAA 781
 |||||
 Db 37 TTTCACCAAAAAAAAAAAAA 19

```

RESULT 2
US-08-377-292-2/c
; Sequence 2, Application US/08377292
; Patent No. 5693615
; GENERAL INFORMATION:
; APPLICANT: STONE, ROGER L.
; TITLE OF INVENTION: THERAPEUTIC FORMULAS FOR OSTEOINDUCTION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Procter & Gamble Company
; STREET: 11810 East Miami River Road
; CITY: Cincinnati
; STATE: Ohio
; COUNTRY: U.S.A.
; ZIP: 45239-8707
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/377,292
; FILING DATE: 23-JAN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/243,435
; FILING DATE:
; APPLICATION NUMBER: US/08/117,367
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Corstanje, Brahm J.
; REGISTRATION NUMBER: 34,804
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 513-245-2858
; TELEFAX: 513-741-3012
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1547 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-377-292-2

Query Match 2.4%; Score 19; DB 1; Length 1547;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 19; Conservative 0; Mismatches 0; Indels

Qy 263 GCCCGAGGAGCTGGAGGGCC 281
      |||||
Db 119 GCCCGAGGAGCTGGAGGGCC 101

RESULT 3
US-07-679-451-1/c
; Sequence 1, Application US/07679451
; Patent No. 5318898
; GENERAL INFORMATION:
; APPLICANT: Israel, David I.
; TITLE OF INVENTION: IMPROVED PRODUCTION OF RECOMBINANT
; TITLE OF INVENTION: BONE-INDUCING PROTEINS
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Legal Affairs, Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02140-2387
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/679,451
; FILING DATE: 19910402
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/179,100
; FILING DATE: 08-APR-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: McDaniel, Patricia A.
; REGISTRATION NUMBER: 33,194
; REFERENCE/DOCKET NUMBER: GI 5180
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-876-1170
; TELEFAX: 617-876-5851
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1607 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; IMMEDIATE SOURCE:
; CLONE: HUMBMPII-CDNA-39
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..355
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 1544..1607
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: US 5,013,649
; FILING DATE: 08-APR-1988
; PUBLICATION DATE: 07-MAY-1991
; US-07-679-451-1
```

```
Query Match 2.4%; Score 19; DB 1; Length 1607;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 263 GGCCGAGGAGTGGAGGGC 281
      |||||||
DB 151 GGCCGAGGAGTGGAGGGC 133
```

```
RESULT 4
US-07-989-847-1/c
; Sequence 1, Application US/07989847
; Patent No. 5866364
; GENERAL INFORMATION:
; APPLICANT: Israel, David
; TITLE OF INVENTION: Recombinant Bone Morphogenetic Protein
; TITLE OF INVENTION: Heterodimers, Compositions and Methods of Use.
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Legal Affairs, Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: MA
; STATE: MA
; COUNTRY: USA
; ZIP: 02140-2387
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Tape
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/989,847
; FILING DATE:
```

```
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Kapinos, Ellen J.
; REGISTRATION NUMBER: 32,245
; REFERENCE/DOCKET NUMBER: GI-5192B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-876-1170
; TELEFAX: 617-876-5851
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1607 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 356..1543
; US-07-989-847-1
```

```
Query Match 2.4%; Score 19; DB 2; Length 1607;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 263 GGCCGAGGAGTGGAGGGC 281
      |||||||
DB 151 GGCCGAGGAGTGGAGGGC 133
```

```
RESULT 5
US-07-721-847A-3/c
; Sequence 3, Application US/07721847A
; Patent No. 6150328
; GENERAL INFORMATION:
; APPLICANT: Wang, Elizabeth A.
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wozney, John M.
; TITLE OF INVENTION: No. 6150328el BMP Products
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LEGAL AFFAIRS, GENETICS INSTITUTE, INC.
; STREET: 87 CAMBRIDGEPARK DRIVE
; CITY: CAMBRIDGE
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/721,847A
; FILING DATE: 14-JUN-1991
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kapinos, Ellen J.
; REGISTRATION NUMBER: 32,245
; REFERENCE/DOCKET NUMBER: 5160C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-876-1170
; TELEFAX: 617-876-5851
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1607 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
```

```
; CELL TYPE: Osteosarcoma Cell Line
; CELL LINE: U-20S
; IMMEDIATE SOURCE:
; LIBRARY: U20S cDNA in Lambda GT10
; CLONE: Lambda U20S-39
; POSITION IN GENOME:
; UNITS: bp
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 356..1546
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1202..1543
; FEATURE:
; NAME/KEY: mRNA
; LOCATION: 14..1607
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 356..424
; US-07-721-847A-3
Query Match 2.4%; Score 19; DB 3; Length 1607;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 263 GCGCGAGGAGTGGAGGGC 281
      |||
DB 151 GCGCGAGGAGTGGAGGGC 133

RESULT 6
US-08-469-411-1/c
; Sequence 1, Application US/08469411
; Patent No. 6190880
; GENERAL INFORMATION:
; APPLICANT: Israel, David
; TITLE OF INVENTION: Recombinant Bone Morphogenetic Protein
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Legal Affairs, Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02140-2387
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Tape
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,411
; FILING DATE: 06-Jun-1995
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Kapinos, Ellen J.
; REGISTRATION NUMBER: 32,245
; REFERENCE/DOCKET NUMBER: GI-5192B-CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-498-8622
; TELEFAX: 617-876-5851
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1607 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 356..1543
```

```
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-469-411-1
Query Match 2.4%; Score 19; DB 4; Length 1607;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 263 GCGCGAGGAGTGGAGGGC 281
      |||
DB 151 GCGCGAGGAGTGGAGGGC 133

RESULT 7
US-08-925-779-3/C
; Sequence 3, Application US/08925779
; Patent No. 6245889
; GENERAL INFORMATION:
; APPLICANT: Wang, Elizabeth A.
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wozney, John M.
; TITLE OF INVENTION: No. 6245889el BMP Products
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LEGAL AFFAIRS, GENETICS INSTITUTE, INC.
; STREET: 87 CAMBRIDGEPARK DRIVE
; CITY: CAMBRIDGE
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/925,779
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/721,847
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kapinos, Ellen J.
; REGISTRATION NUMBER: 32,245
; REFERENCE/DOCKET NUMBER: 5160C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-876-1170
; TELEFAX: 617-876-5851
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1607 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; CELL TYPE: Osteosarcoma Cell Line
; CELL LINE: U-20S
; IMMEDIATE SOURCE:
; LIBRARY: U20S cDNA in Lambda GT10
; CLONE: Lambda U20S-39
; POSITION IN GENOME:
; UNITS: bp
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 356..1546
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1202..1543
```

```
; NAME/KEY: mRNA
; LOCATION: 14..1607
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 356..424
; US-08-925-779-3

Query Match      2.4%; Score 19; DB 4; Length 1607;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 263 GCCCGAGGAGTGGAGGGC 281
      |||||
Db 151 GCCCGAGGAGTGGAGGGC 133

RESULT 8
5166058-3/c
; Patent No. 5166058
; APPLICANT: WANG, ELIZABETH A.; WOZNEY, JOHN M.; RPSN, VICKI A.
; TITLE OF INVENTION: DNA SEQUENCES ENCODING THE OSTEOINDUCTIVE
; PROTEINS
; NUMBER OF SEQUENCES: 19
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/378,537
; FILING DATE: 11-JUL-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 179,100
; FILING DATE: 08-APR-1988
; APPLICATION NUMBER: 28,285
; FILING DATE: 20-MAR-1987
; APPLICATION NUMBER: 943,332
; FILING DATE: 17-DEC-1986
; APPLICATION NUMBER: 880,776
; FILING DATE: 01-JUL-1986
; SEQ ID NO: 3
; LENGTH: 1607
5166058-3

Query Match      2.4%; Score 19; DB 6; Length 1607;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 263 GCCCGAGGAGTGGAGGGC 281
      |||||
Db 151 GCCCGAGGAGTGGAGGGC 133

RESULT 9
US-07-667-276A-3/c
; Sequence 3, Application US/07667276A
; Patent No. 5470971
; GENERAL INFORMATION:
; APPLICANT: Kondo, Keiji
; APPLICANT: Inouye, Masayori
; TITLE OF INVENTION: STRESS-INDUCED PROTEINS, GENES CODING
; THEREFOR, TRANSFORMED CELLS OF ORGANISMS, METHODS AND
; TITLE OF INVENTION: APPLICATIONS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weiser & Associates
; STREET: 230 S. Fifteenth Street, Suite 500
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/667,276A
```

```
; FILING DATE: 11-MAR-1991
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiser, Gerard J.
; REGISTRATION NUMBER: 19,763
; REFERENCE/DOCKET NUMBER: 377.5351P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-875-8383
; TELEFAX: 215-875-8394
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2017 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Saccharomyces cerevisiae
; STRAIN: S288C
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 484..1725
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1
; OTHER INFORMATION: /note= "Base #1 of Sequence No. 5470971 3
; OTHER INFORMATION: corresponds to base 483 of the sequence listed in
; OTHER INFORMATION: Figure 11 of the application"
US-07-667-276A-3

Query Match      2.4%; Score 19; DB 1; Length 2017;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 763 TTTCACCAAAAAAAAAAAAA 781
      |||||
Db 281 TTTCACCAAAAAAAAAAAAA 263

RESULT 10
US-08-811-481-34
; Sequence 34, Application US/08811481
; Patent No. 6300093
; GENERAL INFORMATION:
; APPLICANT: Kindsvogel, Wayne
; APPLICANT: Jelinek, Laura J.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Hagopian, William A.
; APPLICANT: Lagasse, James W.
; TITLE OF INVENTION: ISLET CELL ANTIGEN 1851
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,481
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Lingenfelter, Susan
; REGISTRATION NUMBER: P-41,156
; REFERENCE/DOCKET NUMBER: 95-36
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6675
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2328 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-08-811-481-34

Query Match 2.4%; Score 19; DB 4; Length 2328;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 763 TTCACCAAAAAAAAAAAAAA 781
|||||
Db 2279 TTCACCAAAAAAAAAAAAAA 2297

RESULT 11
US-08-377-687-48
; Sequence 48, Application US/08377687
; Patent No. 5538525
; GENERAL INFORMATION:
; APPLICANT: BROEKAERT, WILLEM F.
; APPLICANT: CAMMUE, BRUNO P.A.
; APPLICANT: OSBORN, RUPERT W.
; APPLICANT: REES, SARAH B.
; APPLICANT: TERRAS, FRANKY R.G.
; APPLICANT: VANDERLEYDEN, JOZEF
; TITLE OF INVENTION: BIOCIDAL PROTEINS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARB Y & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/377,687
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/002,480
; FILING DATE: 04-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 414 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..255
; US-08-777-192-48

Query Match 2.3%; Score 18; DB 1; Length 414;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 764 TTCACCAAAAAAAAAAAAAA 781
|||||
Db 390 TTCACCAAAAAAAAAAAAAA 407

RESULT 12
US-08-777-192-48
; Sequence 48, Application US/08777192
; Patent No. 5824869
; GENERAL INFORMATION:
; APPLICANT: BROEKAERT, WILLEM F.
; APPLICANT: CAMMUE, BRUNO P.A.
; APPLICANT: OSBORN, RUPERT W.
; APPLICANT: REES, SARAH B.
; APPLICANT: TERRAS, FRANKY R.G.
; APPLICANT: VANDERLEYDEN, JOZEF
; TITLE OF INVENTION: BIOCIDAL PROTEINS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARB Y & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/777,192
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/002,480
; FILING DATE: 04-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 414 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..255
; US-08-777-192-48

Query Match 2.3%; Score 18; DB 1; Length 414;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 764 TTCACCAAAAAAAAAAAAAA 781
|||||
Db 390 TTCACCAAAAAAAAAAAAAA 407
```



```

RESULT 13
US-08-971-982-48
: Sequence 48, Application US/08971982
: Patent No. 6187904
: GENERAL INFORMATION:
: APPLICANT: BROEKAERT, WILLEM F.
: CAMMUE, BRUNO P.A.
: OSBORN, RUPERT W.
: REES, SARAH B.
: TERRAS, FRANKY R.G.
: VANDERLEYDEN, JOZEF
: TITLE OF INVENTION: BIOCIDAL PROTEINS
: NUMBER OF SEQUENCES: 59
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CUSHMAN DABY & CUSHMAN
: STREET: 1100 NEW YORK AVENUE, N.W.
: CITY: WASHINGTON
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/971.982
: FILING DATE: 17-No. 6187904-1997
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/002.480
: FILING DATE: 04-JAN-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: KOKULIS, PAUL N.
: REGISTRATION NUMBER: 16,773
: REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-861-3000
: TELEFAX: 202-822-0944
: INFORMATION FOR SEQ ID NO: 48:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 414 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 16..255
: SEQUENCE DESCRIPTION: SEQ ID NO: 48:
US-08-971-982-48
Query Match 2.3%; Score 18; DB 4; Length 414;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 764 TTCACCAAAAAAAAAAAAA 781
Db 390 TTCACCAAAAAAAAAAAAA 407
RESULT 14
US-08-549-846-2
: Sequence 2, Application US/08549846
: Patent No. 6083751
: GENERAL INFORMATION:
: APPLICANT: FELDHAUS, ANDREW L.
: APPLICANT: JONES, LORI A.
: TITLE OF INVENTION: CHIMERIC RECEPTORS FOR THE GENERATION OF
: TITLE OF INVENTION: SELECTIVITY-ACTIVATABLE T(h)-INDEPENDENT CYTOTOXIC T CELLS
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: MORRISON & FOERSTER

```

```

; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5137 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US96-01314-39

Query Match          2.3%; Score 18; DB 5; Length 5137;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 520 CAGCTGAGATCCAGTGC 537
    |||||||
Db 1971 CAGCTGAGATCCAGTGC 1988

RESULT 17
US-08-611-280-4
; Sequence 4, Application US/08611280
; Patent No. 5891666
; GENERAL INFORMATION:
; APPLICANT: Matsuyama, Toshifumi
; APPLICANT: Grossman, Alex
; APPLICANT: Richardson, Christopher D.
; TITLE OF INVENTION: NOVEL GENES ENCODING LSIRF POLYPEPTIDES
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Canada Inc.
; STREET: 6733 Mississauga Road, Suite 303
; CITY: Mississauga
; STATE: Ontario
; COUNTRY: Canada
; ZIP: L5N 6J8
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,280
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oleski, Nancy A.
; REGISTRATION NUMBER: 34,688
; REFERENCE/DOCKET NUMBER: A-338A
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12537 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-611-280-4

Query Match          2.3%; Score 18; DB 2; Length 12537;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 492 GGCAGCAGACAGAGAA 509
    |||||||
Db 4358 GGCAGCAGACAGAGAA 4375

RESULT 18
US-09-195-940-4
; Sequence 4, Application US/09195940
; Patent No. 6258935
; GENERAL INFORMATION:
; APPLICANT: Matsuyama, Toshifumi
; APPLICANT: Grossman, Alex
; APPLICANT: Richardson, Christopher D.
; TITLE OF INVENTION: NOVEL GENES ENCODING LSIRF POLYPEPTIDES
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Canada Inc.
; STREET: 6733 Mississauga Road, Suite 303
; CITY: Mississauga
; STATE: Ontario
; COUNTRY: Canada
; ZIP: L5N 6J8
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/195,940
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/611,280
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Oleski, Nancy A.
; REGISTRATION NUMBER: 34,688
; REFERENCE/DOCKET NUMBER: A-338A
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12537 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-195-940-4

Query Match      2.3%; Score 18; DB 4; Length 12537;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 492 GGGCAGCAGACAGAGAA 509
      |||
DB 4358 GGGCAGCAGACAGAGAA 4375

RESULT 19
US-08-955-937A-3
; Sequence 3, Application US/08955937A
; Patent No. 6020161
; GENERAL INFORMATION:
; APPLICANT: WU, SHUJIAN
; APPLICANT: SWEET, RAYMOND
; APPLICANT: TRUNEH, ALEMESEGED
; TITLE OF INVENTION: PIGR-1, A MEMBER OF IMMUNOGLOBULIN
; TITLE OF INVENTION: GENE SUPERFAMILY
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/955,937A
; FILING DATE: 17-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/056,152
; FILING DATE: 19-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70228
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 708 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-955-937A-3

Query Match      2.2%; Score 17; DB 3; Length 708;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 273 TGGAGGGGCTCAGGCGA 289
      |||
DB 382 TGGAGGGGCTCAGGCGA 398

RESULT 20
US-09-300-985-3
; Sequence 3, Application US/09300985A
; Patent No. 6232441
; GENERAL INFORMATION:
; APPLICANT: WU, SHUJIAN
; APPLICANT: SWEET, RAYMOND W.
; APPLICANT: TRUNEH, ALEMESEGED
; APPLICANT: HURLE, MARK ROBERT
; TITLE OF INVENTION: PIGR-1, A MEMBER OF IMMUNOGLOBULIN GENE
; FILE REFERENCE: GH-70228-1
; CURRENT APPLICATION NUMBER: US/09/300,985A
; CURRENT FILING DATE: 1999-04-28
; EARLIER APPLICATION NUMBER: US 08/955,937
; EARLIER FILING DATE: 1997-10-22
; EARLIER APPLICATION NUMBER: US 60/056,152
; EARLIER FILING DATE: 1997-08-19
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 708
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (475)(620)(660)
; US-09-300-985-3

Query Match      2.2%; Score 17; DB 4; Length 708;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 273 TGGAGGGGCTCAGGCGA 289
      |||
DB 382 TGGAGGGGCTCAGGCGA 398

RESULT 21
US-08-123-161A-11
; Sequence 11, Application US/08123161A
; Patent No. 5449616
; GENERAL INFORMATION:
; APPLICANT: Campbell, Kevin P.
; APPLICANT: Roberts, Steven L.
; APPLICANT: Anderson, Richard D.
; APPLICANT: Ibraghimov, Oxana B.
; APPLICANT: Yang, Bin
; TITLE OF INVENTION: NUCLEIC ACID ENCODING DYSTROPHIN-ASSOCIATED
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kevin M. Farrell, P.C.
; STREET: P.O. Box 999
; CITY: York Harbor
; STATE: ME
; COUNTRY: USA

```

```
;
;
; ZIP: 03911
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/123,161A
; FILING DATE: 16-SEP-93
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/946,234
; FILING DATE: 14-SEP-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Farrell, Kevin M.
; REGISTRATION NUMBER: 35,505
; TELEPHONE: (207) 363-0558
; TELEFAX: (207) 363-0528
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1396 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4..1164
;
US-08-123-161A-11

Query Match 2.2%; Score 17; DB 1; Length 1396;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 765 TCACCAAAAAAAAAAAAA 781
Db 1366 TCACCAAAAAAAAAAAAA 1382

RESULT 22
US-08-483-278-11
; Sequence 11, Application US/08483278
; Patent No. 5686073
; GENERAL INFORMATION:
; APPLICANT: Campbell, Kevin P.
; APPLICANT: Ibraghimov, Oxana B.
; APPLICANT: Ervasti, James M.
; APPLICANT: Levellie, Cynthia J.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING DYSTROPHIN-ASSOCIATED
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kevin M. Farrell, P.C.
; STREET: P.O. Box 999
; CITY: York Harbor
; STATE: ME
; COUNTRY: USA
; ZIP: 03911
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,278
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/123,161
; FILING DATE: 16-SEP-93
; ATTORNEY/AGENT INFORMATION:
```

```
;
;
; NAME: Farrell, Kevin M.
; REGISTRATION NUMBER: 35,505
; REFERENCE/DOCKET NUMBER: UIRF89-11A5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (207) 363-0558
; TELEFAX: (207) 363-0528
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1396 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4..1164
;
US-08-483-278-11

Query Match 2.2%; Score 17; DB 1; Length 1396;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 765 TCACCAAAAAAAAAAAAA 781
Db 1366 TCACCAAAAAAAAAAAAA 1382

RESULT 23
US-09-189-760-5
; Sequence 5, Application US/09189760
; Patent No. 6031078
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-046CP2
; CURRENT APPLICATION NUMBER: US/09/189,760
; CURRENT FILING DATE: 1998-11-10
; EARLIER APPLICATION NUMBER: 09/163,116
; EARLIER FILING DATE: 1998-09-29
; EARLIER APPLICATION NUMBER: 60/089,467
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: (PENDING)
; EARLIER FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1529
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(749)
;
US-09-189-760-5

Query Match 2.2%; Score 17; DB 3; Length 1529;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 765 TCACCAAAAAAAAAAAAA 781
Db 1420 TCACCAAAAAAAAAAAAA 1436

RESULT 24
US-09-188-811-5
; Sequence 5, Application US/09188811
; Patent No. 6037148
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-046CP
```

```
; CURRENT APPLICATION NUMBER: US/09/188,811
; CURRENT FILING DATE: 1998-11-09
; EARLIER APPLICATION NUMBER: 09/163,116
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1529
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(749)
US-09-188-811-5

Query Match          2.2%; Score 17; DB 3; Length 1529;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 765 TCACCAAAAAAAAAAAAA 781
Db 1420 TCACCAAAAAAAAAAAAA 1436

RESULT 25
US-09-514-422-5
; Sequence 5, Application US/09514422
; Patent No. 6291193
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: MNI-046CP2
; CURRENT APPLICATION NUMBER: US/09/514,422
; CURRENT FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: US/09/189,760
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 09/163,116
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/089,467
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: (PENDING)
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1529
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(749)
US-09-514-422-5

Query Match          2.2%; Score 17; DB 4; Length 1529;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 765 TCACCAAAAAAAAAAAAA 781
Db 1420 TCACCAAAAAAAAAAAAA 1436

RESULT 26
US-08-749-289-2/c
; Sequence 2, Application US/08749289
; Patent No. 5955301
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Coleman, Roger
; TITLE OF INVENTION: NOVEL HUMAN GLUTAMATE-BINDING
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 3
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/749,289
; FILING DATE: Filed Herewith
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0160 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1770 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: 385116
US-08-749-289-2

Query Match          2.2%; Score 17; DB 2; Length 1770;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 655 GAACGAGTGAAGAGCA 671
Db 1029 GAACGAGTGAAGAGCA 1013

RESULT 27
US-08-878-563A-2
; Sequence 2, Application US/08878563A
; Patent No. 5891674
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: INSULIN RECEPTOR TYROSINE KINASE SUBSTRATE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/878,563A
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
```

```

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0323 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-845-4166
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2080 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRSTNOT04
; CLONE: 918158
;
US-08-878-563A-2
;
Query Match 2.2%; Score 17; DB 2; Length 2080;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 GTCCACAGGCAGATCCA 93
Db 339 GTCCACAGGCAGATCCA 355

RESULT 28
US-09-270-117-2
; Sequence 2, Application US/09270117
; Patent No. 6265550
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: INSULIN RECEPTOR TYROSINE KINASE SUBSTRATE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
;
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/270,117
FILING DATE:
PRIOR APPLICATION DATA:
ATTORNEY/AGENT INFORMATION:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0323 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-845-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2080 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRSTNOT04
CLONE: 918158
;
US-09-270-117-2

```

```

Query Match 2.2%; Score 17; DB 4; Length 2080;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 GTCCACAGGCAGATCCA 93
Db 339 GTCCACAGGCAGATCCA 355

RESULT 29
US-08-955-937A-1
; Sequence 1, Application US/08955937A
; Patent No. 6020161
; GENERAL INFORMATION:
; APPLICANT: WU, SHUJIAN
; APPLICANT: SWEET, RAYMOND
; APPLICANT: TRUNER, ALEMESEGED
; TITLE OF INVENTION: PIGR-1, A MEMBER OF IMMUNOGLOBULIN
; TITLE OF INVENTION: GENE SUPERFAMILY
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RAYNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
;
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/955,937A
FILING DATE: 17-OCT-1997
PRIOR APPLICATION DATA:
CLASSIFICATION: 435
APPLICATION NUMBER: 60/056,152
FILING DATE: 19-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F.
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70228
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2345 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-955-937A-1

Query Match 2.2%; Score 17; DB 3; Length 2345;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 273 TGGAGGGGCTCAGCGA 289
Db 403 TGGAGGGGCTCAGCGA 419

RESULT 30
US-09-300-985-1
; Sequence 1, Application US/09300985A
; Patent No. 6232441
; GENERAL INFORMATION:
; APPLICANT: WU, SHUJIAN
; APPLICANT: SWEET, RAYMOND W.

```

```

; APPLICANT: TRUNEH, ALEMSEGED
; APPLICANT: HURLE, MARK ROBERT
; TITLE OF INVENTION: PIGR-1, A MEMBER OF IMMUNOGLOBULIN GENE
; TITLE OF INVENTION: SUPERFAMILY
; FILE REFERENCE: GH-70228-1
; CURRENT APPLICATION NUMBER: US/09/300,985A
; CURRENT FILING DATE: 1999-04-28
; EARLIER APPLICATION NUMBER: US 08/955,937
; EARLIER FILING DATE: 1997-10-22
; EARLIER APPLICATION NUMBER: US 60/056,152
; EARLIER FILING DATE: 1997-08-19
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2345
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
; US-09-300-985-1

Query Match      2.2%; Score 17; DB 4; Length 2345;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 273 TGGAGGGGCTCAGCGCA 289
      |||||
Db 403 TGGAGGGGCTCAGCGCA 419

RESULT 31
US-09-189-760-1
; Sequence 1, Application US/09189760
; Patent No. 6031078
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: MNI-046CP2
; CURRENT APPLICATION NUMBER: US/09/189,760
; CURRENT FILING DATE: 1998-11-10
; EARLIER APPLICATION NUMBER: 09/163,116
; EARLIER FILING DATE: 1998-09-29
; EARLIER APPLICATION NUMBER: 60/089,467
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: (PENDING)
; EARLIER FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2494
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (164)..(1714)
; US-09-189-760-1

Query Match      2.2%; Score 17; DB 3; Length 2494;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 765 TCACCAAAAAAAAAAAAA 781
      |||||
Db 2385 TCACCAAAAAAAAAAAAA 2401

RESULT 32
US-09-514-422-1
; Sequence 1, Application US/09514422
; Patent No. 6291193
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: MNI-046CP2
; CURRENT APPLICATION NUMBER: US/09/514,422
; CURRENT FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: US/09/189,760
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 09/163,116
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/089,467
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: (PENDING)
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2494
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (164)..(1714)
; US-09-514-422-1

Query Match      2.2%; Score 17; DB 4; Length 2494;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 765 TCACCAAAAAAAAAAAAA 781
      |||||
Db 2385 TCACCAAAAAAAAAAAAA 2401

RESULT 33
US-08-162-081B-35
; Sequence 35, Application US/08162081B
; Patent No. 5824492
; GENERAL INFORMATION:
; APPLICANT: Hiles, Ian Donald; Fry, Michael John; Dhand, Ritu
; APPLICANT: Bala; Waterfield, Michael Derek; Parker, Peter
; APPLICANT: Joseph; Otsu, Masayuki; Panayotou, George; Volinia,
; APPLICANT: Stefano; Gout, Ivan Tarasovitch
; TITLE OF INVENTION: POLYPEPTIDES HAVING KINASE ACTIVITY,
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/162,081B
; FILING DATE: February 7, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB93/00761
; FILING DATE: 13 April 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Pasqualini, Patricia A.
; REGISTRATION NUMBER: 34,894
; REFERENCE/DOCKET NUMBER: LUD 5256
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3207 base pairs
; TYPE: nucleic acid
```

```
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-162-081B-35

Query Match          2.2%  Score 17; DB 1; Length 3207;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 629 CTAAGAGGAACTGT 645
Db 1244 CTAAGAGGAACTGT 1260
|||||

RESULT 34
US-08-780-872-35
; Sequence 35, Application US/08780872
; Patent No. 5846824
; GENERAL INFORMATION:
; APPLICANT: Hiles, Ian Donald; Fry, Michael John; Dhand, Ritu
; APPLICANT: Bala; Waterfield, Michael Derek; Parker, Peter
; APPLICANT: Joseph; Otsu, Masayuki; Panayotou, George; Volinia,
; APPLICANT: Stefano; Gout, Ivan Tarasovitch
; TITLE OF INVENTION: POLYPEPTIDES HAVING KINASE ACTIVITY,
; TITLE OF INVENTION: THEIR PREPARATION AND USE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/085,957
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/780,872
; FILING DATE: 09-JAN-1997
; APPLICATION NUMBER: 08/162,081
; FILING DATE: February 7, 1994
; APPLICATION NUMBER: PCT/GB93/00761
; FILING DATE: 13 April 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Pasqualini, Patricia A.
; REGISTRATION NUMBER: 34,894
; REFERENCE/DOCKET NUMBER: LUD 5256
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3207 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-085-957-35

Query Match          2.2%  Score 17; DB 4; Length 3207;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 629 CTAAGAGGAACTGT 645
Db 1244 CTAAGAGGAACTGT 1260
|||||

RESULT 36
US-08-162-081B-34
; Sequence 34, Application US/08162081B
; Patent No. 5824492
; GENERAL INFORMATION:
; APPLICANT: Hiles, Ian Donald; Fry, Michael John; Dhand, Ritu
; APPLICANT: Bala; Waterfield, Michael Derek; Parker, Peter
; APPLICANT: Joseph; Otsu, Masayuki; Panayotou, George; Volinia,
; APPLICANT: Stefano; Gout, Ivan Tarasovitch
; TITLE OF INVENTION: POLYPEPTIDES HAVING KINASE ACTIVITY,
; TITLE OF INVENTION: THEIR PREPARATION AND USE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
```


; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/162,081B
; FILING DATE: February 7, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB93/00761
; FILING DATE: 13 April 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Pasqualini, Patricia A.
; REGISTRATION NUMBER: 34,894
; TELEPHONE: (212) 838-3884
; TELEFAX: (212) 838-9200
; TELECOMMUNICATION INFORMATION:
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-162-081B-34

Query Match 2.2%; Score 17; DB 1; Length 3240;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 629 CTAAGAGGAACTGT 645
|||||
DB 1244 CTAAGAGGAACTGT 1260

RESULT 37
US-08-780-872-34
; Sequence 34, Application US/08780872
; Patent No. 5848824
; GENERAL INFORMATION:
; APPLICANT: Hiles, Ian Donald; Fry, Michael John; Dhand, Ritu
; APPLICANT: Bala, Waterfield, Michael Derek; Parker, Peter
; APPLICANT: Joseph; Otsu, Masayuki; Panayotou, George; Volinia,
; APPLICANT: Stefano; Gout, Ivan Tarasovitch
; TITLE OF INVENTION: POLYPEPTIDES HAVING KINASE ACTIVITY,
; THEIR PREPARATION AND USE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/780,872
; FILING DATE: 13 April 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Pasqualini, Patricia A.
; REGISTRATION NUMBER: 34,894
; TELEPHONE: (212) 838-3884
; TELEFAX: (212) 838-9200
; TELECOMMUNICATION INFORMATION:
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-085-957-34

Query Match 2.2%; Score 17; DB 4; Length 3240;

Best Local Similarity 100.0%; Pred. No. 46;
Matches 17: Conservative 0; Mismatches

Qy	629	CTAAGAGGAACACTGT	645
Db	1244	CTAAGAGGAACACTGT	1260

RESULT 39

US-08-162-081B-32
; Sequence 32, Application US/08162081B
; Patent No. 5824492

```

GENERAL INFORMATION:
APPLICANT: Hiles, Ian Donald; Fry, Michael John; Dhand, Ritu
APPLICANT: Bala; Waterfield, Michael Derek; Parker, Peter
APPLICANT: Joseph; Otsu, Masayuki; Panayotou, George; Volinia;
APPLICANT: Stefano; Gout, Ivan Tarasovitch
TITLE OF INVENTION: POLYPEPTIDES HAVING KINASE ACTIVITY,
TITLE OF INVENTION: THEIR PREPARATION AND USE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022

```

```

; ZIF. 10022
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
;

```

```

;
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect

```

/ SOFTWARE: MICROPERFECT
 /
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/162,081B
 / FILING DATE: February 7, 1994
 /

; FILING DATE: FEBRUARY 1968
 ; CLASSIFICATION: 435

CLASSIFICATION: 425
PRIOR APPLICATION DATA: PCT/GB93/00761
APPLICATION NUMBER: 13 April 1993
FILING DATE: 13 April 1993

ATTORNEY/AGENT INFORMATION:
NAME: Pasqualini, Patricia A.

REGISTRATION NUMBER: 34,894
REFERENCE/DOCKET NUMBER: LUD 5256
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884

; TELFAX. (412) 650-3004
; INFORMATION FOR SEQ ID NO: 32:

; SEQUENCE CHARACTERISTICS:

```

; LENGTH: 3412 base pairs
;
; TYPE: nucleic acid
;
; STRANDEDNESS: single or double
;
; TOPOLOGY: linear
;

```

```
;
;
FEATURE:
```

NAME/KEY: CDS

; LOCATION: 1..3204

```

; OTHER INFORMATION: /standard_name="CDS
IIS-08-162-081B-32

```

US-08-162-081B-32

Query Match 2.2%; Score 17; DB 1; Length 3412;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 17; Conservative 0; Mismatches 0; Indels

Qy	629	CTAAAGAGGAACACTGT	645
Db	1244	CTAAAGAGGAACACTGT	1260

RESULT 40

US-08-780-872-32

; Sequence 32, Application US/08780872
; Patent No. 5846824

GENERAL INFORMATION:

; APPLICANT: Hiles, Ian Donald; Fry, Michael John; Dhand, Ritu

**ORIGINAL
COPY**

91	21	2.7	498	10	BF045827	BF045827 BP250005A	C 164	19	2.4	251	9	AI880893	AI880893 at19f03.x
c 92	21	2.7	500	9	AI402186	AI402186 GH05786.3	c 165	19	2.4	254	10	BI428179	BI428179 fr75e08.x
93	21	2.7	525	10	BM419880	BM419880 RQ20B01.O	c 166	19	2.4	257	10	EG220980	EG220980 RST40780
c 94	21	2.7	547	10	BE681607	BE681607 I79329.MA	c 167	19	2.4	257	10	BG579452	BG579452 dab75b05.
c 95	21	2.7	564	12	BH616877	BH616877 BMBAC310H	c 168	19	2.4	261	10	BF466628	BF466628 UI-M-CG0p
c 96	21	2.7	886	10	BI695973	BI695973 603346165	c 169	19	2.4	262	9	AW066672	AW066672 683006A02
c 97	21	2.7	923	12	CNS003XB	AL065285 Drosophil	c 170	19	2.4	262	10	BM334876	BM334876 MEST143-C
c 98	21	2.7	1527	12	AG135591	AG135591 Pap trogl	c 171	19	2.4	262	10	BE995399	BE995399 UI-M-CG0p
c 99	20	2.6	223	10	BF470674	BF470674 UI-M-BH3-	c 172	19	2.4	265	9	AV742275	AV742275 AV742275
c 100	20	2.6	240	9	AI556212	AI556212 UI-R-C2p-	c 173	19	2.4	276	12	AQ108362	AQ108362 C1T-HSP-2
c 101	20	2.6	277	10	BE949311	BE949311 UI-M-BH3-	c 174	19	2.4	278	12	AZ660900	AZ660900 IM0539A16
c 102	20	2.6	292	10	BE990047	BE990047 UI-M-BZ1-	c 175	19	2.4	284	9	AU031091	AU031091 AU031091
c 103	20	2.6	312	10	BF408985	BF408985 UI-R-BJ2-	c 176	19	2.4	285	9	BB420880	BB420880 BB420880
c 104	20	2.6	343	9	AW204002	AW204002 UI-H-B11-	c 177	19	2.4	287	10	BF711597	BF711597 MI-P-A1-a
c 105	20	2.6	343	10	BF398312	BF398312 UI-R-BS2-	c 178	19	2.4	289	9	BB251667	BB251667 BB251667
c 106	20	2.6	351	10	BE525130	BE525130 M58H1STM	c 179	19	2.4	293	9	BB009833	BB009833 BB009833
c 107	20	2.6	367	10	BE995270	BE995270 UI-M-CG0p	c 180	19	2.4	298	9	AW655358	AW655358 105984.MA
c 108	20	2.6	379	10	D55484	D55484 HUM180B09B	c 181	19	2.4	308	9	AW523054	AW523054 UI-R-B00-
c 109	20	2.6	386	9	AA817964	AA817964 UI-R-A0-a	c 182	19	2.4	308	10	BG555746	BG555746 df20h05.x
c 110	20	2.6	398	9	AU1715284	AU1715284 UI-R-Y0-a	c 183	19	2.4	313	9	BB396707	BB396707 BB396707
c 111	20	2.6	442	9	AU172665	AU172665 AU172665	c 184	19	2.4	316	9	BB451464	BB451464 BB451464
c 112	20	2.6	456	9	AV739448	AV739448 AV739448	c 185	19	2.4	318	9	AW073875	AW073875 xb04c10.x
c 113	20	2.6	465	9	BB855812	BB855812 BB855812	c 186	19	2.4	323	10	BM372304	BM372304 EBR003.SQ
c 114	20	2.6	466	12	AZ294141	AZ294141 RPCI-23-1	c 187	19	2.4	325	9	AI510997	AI510997 UI-R-C2p-
c 115	20	2.6	467	9	AW159644	AW159644 zb05b07.x	c 188	19	2.4	325	10	BM337332	BM337332 MEST149-C
c 116	20	2.6	474	9	AW400088	AW400088 707051B11	c 189	19	2.4	326	9	AA179328	AA179328 zp12b02.r
c 117	20	2.6	479	9	AW452932	AW452932 UI-H-BW1-	c 190	19	2.4	327	4	BB707581	BB707581 Mus muscu
c 118	20	2.6	510	10	BI400173	BI400173 MI-P-AV1-	c 191	19	2.4	334	10	BI491308	BI491308 df06h11.w
c 119	20	2.6	519	10	BM266318	BM266318 VL83.CDNA	c 192	19	2.4	343	9	AA497170	AA497170 fa01g08.s
c 120	20	2.6	548	9	BE113957	BE113957 UI-R-BJ1-	c 193	19	2.4	348	10	BI000385	BI000385 MR3-HN006
c 121	20	2.6	570	12	AZ467193	AZ467193 IM0278J20	c 194	19	2.4	349	9	AI304985	AI304985 EST00090
c 122	20	2.6	576	12	CNS03TVM	AL260347 Tetraodon	c 195	19	2.4	349	9	BE138496	BE138496 xr76a10.x
c 123	20	2.6	605	12	AQ959104	AQ959104 LEREH33TF	c 196	19	2.4	350	10	BF273284	BF273284 GA_Eb001
c 124	20	2.6	613	12	AZ821669	AZ821669 ZM0094102	c 197	19	2.4	351	9	AI544994	AI544994 fb73d03.x
c 125	20	2.6	646	10	BG841415	BG841415 MEST21-D0	c 198	19	2.4	356	9	AU101146	AU101146 AU101146
c 126	20	2.6	670	12	CNS02F51	AL194590 Tetraodon	c 199	19	2.4	356	10	BI493375	BI493375 df100C06.
c 127	20	2.6	693	12	AZ461440	AZ461440 IM0267L15	c 200	19	2.4	360	9	A1029102	A1029102 UI-R-C0-1
c 128	20	2.6	714	12	AQ959105	AQ959105 LEREH33TF	c 201	19	2.4	360	10	C40203	C40203 C40203.YuJ1
c 129	20	2.6	776	12	BH054219	BH054219 RPCI-24-3	c 202	19	2.4	363	9	AI396151	AI396151 487012A09
c 130	20	2.6	836	10	BI102740	BI102740 602888317	c 203	19	2.4	363	9	AI396189	AI396189 487012F08
c 131	20	2.6	884	10	BF573878	BF573878 602132085	c 204	19	2.4	365	9	AI396159	AI396159 487012B08
c 132	20	2.6	898	12	CNS0123J	AL101065 Drosophil	c 205	19	2.4	366	10	BF597638	BF597638 su87d08.y
c 133	20	2.6	992	10	BG391711	BG391711 602417991	c 206	19	2.4	368	10	BG294675	BG294675 602392041
c 134	20	2.6	1019	12	CNS03XRS	AL265370 Tetraodon	c 207	19	2.4	368	10	BI806960	BI806960 S080B10.S
c 135	20	2.6	1135	12	AG126408	AG126408 Pan trogl	c 208	19	2.4	369	9	AA271000	AA271000 va85d07.r
c 136	20	2.6	2087	11	BC010625	BC010625 Homo sapi	c 209	19	2.4	371	9	AA937811	AA937811 nw89b06.s
c 137	19	2.4	101	9	BE043556	BE043556 hk40b05.y	c 210	19	2.4	371	9	AI395271	AI395271 MA002782.
c 138	19	2.4	127	9	AW881391	AW881391 QV4-OT003	c 211	19	2.4	372	10	BI900766	BI900766 ib81a08.y
c 139	19	2.4	138	10	BE663162	BE663162 133617.MA	c 212	19	2.4	374	10	F36051	F36051 HSPD33359.H
c 140	19	2.4	150	9	AI940262	AI940262 CM3-CT003	c 213	19	2.4	378	9	AI974482	AI974482 T110432e
c 141	19	2.4	160	9	BE111941	BE111941 UI-R-BS1-	c 214	19	2.4	378	12	AQ731336	AQ731336 HS-5512_B
c 142	19	2.4	168	10	N57994	N57994 yv62f07.s1	c 215	19	2.4	385	9	BE104509	BE104509 UI-R-BX0-
c 143	19	2.4	169	9	AI934411	AI934411 wp05d12.x	c 216	19	2.4	387	10	BG441767	BG441767 GA_Ea001
c 144	19	2.4	169	10	BG442916	BG442916 GA_Ea001	c 217	19	2.4	388	10	BM358944	BM358944 GA_Ea001
c 145	19	2.4	174	9	AV140319	AV140319 AV140319	c 218	19	2.4	388	12	AQ526468	AQ526468 HS_5309.B
c 146	19	2.4	176	9	A1828599	AI828599 w142g04.x	c 219	19	2.4	391	9	BE103967	BE103967 UI-R-BX0-
c 147	19	2.4	180	10	BF713062	BF713062 MI-P-H3-a	c 220	19	2.4	396	10	BG195195	BG195195 RST14374
c 148	19	2.4	189	10	BI294600	BI294600 '11-R-DK0-	c 221	19	2.4	397	9	AA086008	AA086008 zn64h03.r
c 149	19	2.4	190	10	BI805139	BI805139 S005E11.S	c 222	19	2.4	398	9	AI396158	AI396158 487012B07
c 150	19	2.4	191	9	AU071425	AU071425 AU071425	c 223	19	2.4	399	10	BG195736	BG195736 RST14930
c 151	19	2.4	209	9	AW834597	AW834597 MR2-TT001	c 224	19	2.4	400	10	BG994061	BG994061 PMO-HT091
c 152	19	2.4	209	9	BE149111	BE149111 RC4-HT025	c 225	19	2.4	402	10	BM075692	BM075692 MEST359-B
c 153	19	2.4	215	9	AI585048	AI585048 fb69g06.x	c 226	19	2.4	404	9	AI582900	AI582900 ts07d04.x
c 154	19	2.4	220	10	BF820389	BF820389 MRI-RT004	c 227	19	2.4	404	9	AV706881	AV706881 AV706881
c 155	19	2.4	221	9	AI563608	AI563608 vx91f04.x	c 228	19	2.4	405	10	BG895250	BG895250 358571.MA
c 156	19	2.4	228	10	BG272561	BG272561 nah34908.	c 229	19	2.4	407	9	AI454462	AI454462 UI-R-C2p-
c 157	19	2.4	229	9	AW600647	AW600647 707104E02	c 230	19	2.4	411	9	AU183353	AU183353 AU183353
c 158	19	2.4	236	9	BB048238	BB048238 BB048238	c 231	19	2.4	412	9	BE232382	BE232382 137390.MA
c 159	19	2.4	239	9	AI917241	AI917241 ts53d09.x	c 232	19	2.4	413	10	BE984741	BE984741 UI-M-CG0p
c 160	19	2.4	240	9	AU071424	AU071424 AU071424	c 233	19	2.4	413	10	BE984744	BE984744 UI-M-CG0p
c 161	19	2.4	242	10	BI004685	BI004685 CMO-HN020	c 234	19	2.4	416	12	AZ135172	AZ135172 OSUNB0011
c 162	19	2.4	248	9	BE216655	BE216655 HV_CEB001	c 235	19	2.4	417	10	BI290953	BI290953 UI-R-DK0-
163	19	2.4	249	9	AI588892	AI588892 tq28h09.x	c 236	19	2.4	418	9	AA252078	AA252078 zr63f04.s

c 237	19	2.4	420	10	B8811889	B8811889 PMO-AN003	310	19	2.4	541	12	AZ738381	AZ738381 RPC1-24-1
c 238	19	2.4	421	12	AQ277358	CITBI-EI-	c 311	19	2.4	544	9	AM049843	AM049843 UI-M-BH1-
c 239	19	2.4	422	9	A1662312	ms08c10.x	312	19	2.4	544	10	BG756327	BG756327 602713646
c 240	19	2.4	423	9	BB699421	BB699421	313	19	2.4	544	10	B1795438	B1795438 602106 E
c 241	19	2.4	426	10	BF516028	BF516028 UI-H-BW1-	314	19	2.4	546	10	BG763052	BG763052 602734931
c 242	19	2.4	427	9	AW562844	660067F01	315	19	2.4	546	10	BG769167	BG769167 602743251
c 243	19	2.4	431	9	AI442274	sa26c11.x	316	19	2.4	547	9	BE234412	BE234412 141414 MA
c 244	19	2.4	434	9	AI713902	UI-R-AC1-	c 317	19	2.4	547	10	BM332924	BM332924 MEST181-2
c 245	19	2.4	435	9	AA118985	mp64005.f	c 318	19	2.4	548	10	BM266863	BM266863 MEST387-P
c 246	19	2.4	435	9	AW135797	UI-H-B11-	c 319	19	2.4	548	12	AZ969771	AZ969771 2M0242101
c 247	19	2.4	435	10	BB848045	uw32b06.y	320	19	2.4	549	10	BG684619	BG684619 602636393
c 248	19	2.4	438	9	AW973753	EST385854	c 321	19	2.4	551	10	BM268493	BM268493 MEST396-D
c 249	19	2.4	443	10	BM333167	MEST184-H	322	19	2.4	551	12	AZ394429	AZ394429 1M0158E01
c 250	19	2.4	448	9	AU090156	AU090156	323	19	2.4	554	10	B1860057	B1860057 603387235
c 251	19	2.4	450	12	AQ144755	HS_3092_A	c 324	19	2.4	556	10	BG194648	BG194648 RST13812
c 252	19	2.4	451	9	AV618633	AV618633	325	19	2.4	558	10	BG204372	BG204372 RST23774
c 253	19	2.4	458	10	BF410519	UI-R-CA0-	326	19	2.4	558	10	BM421785	BM421785 V019G10 O
c 254	19	2.4	459	9	BB863368	BB863368	327	19	2.4	559	9	AV738206	AV738206 AV738206
c 255	19	2.4	462	9	BE111663	UI-R-BJ1-	328	19	2.4	559	10	BE500751	BE500751 MHE0991-O
c 256	19	2.4	464	10	BB853234	uw86d09.x	329	19	2.4	561	10	BE410489	BE410489 601301913
c 257	19	2.4	466	9	AW161946	au71h09.x	c 330	19	2.4	562	9	AW180963	AW180963 MGA0095r
c 258	19	2.4	466	10	BM338233	MEST223-F	c 331	19	2.4	567	10	BG266572	BG266572 1000097D1
c 259	19	2.4	468	10	BM331984	MEST173-E	c 332	19	2.4	567	10	BM351647	BM351647 MEST344-C
c 260	19	2.4	469	12	AQ607665	HS_5408_A	c 333	19	2.4	569	12	AZ477423	AZ477423 1M0296H23
c 261	19	2.4	471	12	AQ887433	HS_5555_A	c 334	19	2.4	571	9	BB652116	BB652116 606036808
c 262	19	2.4	478	9	AI560424	tn09h12.x	c 335	19	2.4	574	9	A1834591	A1834591 606070B09
c 263	19	2.4	478	9	AL385706	MCBC30B05	c 336	19	2.4	575	9	AI734767	AI734767 606036808
c 264	19	2.4	479	10	BG375691	UI-R-CS0-	c 337	19	2.4	575	10	B1865772	B1865772 ft26h07.x
c 265	19	2.4	479	10	BI272150	NF021C11F	c 338	19	2.4	575	10	BE853468	BE853468 uw32b06.x
c 266	19	2.4	480	9	AW503991	UI-HF-BN0	c 339	19	2.4	585	10	BM421924	BM421924 V021H11 O
c 267	19	2.4	480	9	BB635613	BB635613	c 340	19	2.4	587	9	AW700855	AW700855 p845a02.y
c 268	19	2.4	481	9	AU095220	AU095220	c 341	19	2.4	593	9	AV760729	AV760729 AV760729
c 269	19	2.4	481	10	BI495116	BI495116 dfl15b09.	c 342	19	2.4	593	10	BJ025715	BJ025715 BJ025715
c 270	19	2.4	486	9	AA075354	zm87q07.s	343	19	2.4	593	12	AZ744550	AZ744550 RPC1-24-1
c 271	19	2.4	486	10	BI788550	BI788550 ie40e07.x	344	19	2.4	593	12	AQ381037	AQ381037 RPC111-13
c 272	19	2.4	487	9	AW531279	UI-R-RS0-	c 345	19	2.4	597	10	BM266342	BM266342 MEST380-B
c 273	19	2.4	488	10	BM419343	ROI2E11 O	c 346	19	2.4	597	12	AQ381081	AQ381081 RPC111-13
c 274	19	2.4	490	10	BE333018	us97d06.y	c 347	19	2.4	610	10	BM333589	BM333589 MEST158-A
c 275	19	2.4	491	9	AA806538	AA806538 oc24b02.s	c 348	19	2.4	610	12	AZ763823	AZ763823 1M0559023
c 276	19	2.4	491	9	AA526015	nl58d06.s	c 349	19	2.4	611	10	BG344727	BG344727 MEST399-D
c 277	19	2.4	492	9	AA530910	AA530910 nl97f11.s	c 350	19	2.4	615	10	BM3337601	BM3337601 MEST209-A
c 278	19	2.4	492	10	BF278981	GA_Eb003	c 351	19	2.4	617	12	AQ256761	AQ256761 nbxb00640
c 279	19	2.4	494	10	BF350609	BF350609 PMI-HT034	c 352	19	2.4	617	12	AQ365641	AQ365641 nbxb00640
c 280	19	2.4	498	10	BM351415	BM351415 MEST341-B	c 353	19	2.4	617	12	AQ741820	AQ741820 HS_5568_B
c 281	19	2.4	503	10	BM332858	BM332858 MEST180-E	c 354	19	2.4	621	9	AA596830	AA596830 BB659830
c 282	19	2.4	503	10	BF177687	BF177687 Ljirnpet	c 355	19	2.4	626	9	AA968363	AA968363 uc71f02.x
c 283	19	2.4	509	9	AW424169	AW424169 sh61g12.y	c 356	19	2.4	630	10	BM267162	BM267162 MEST391-A
c 284	19	2.4	509	9	AW424169	AW424169 sh61g12.y	c 357	19	2.4	631	9	AL638174	AL638174 AL638174
c 285	19	2.4	510	10	BG397853	BG397853 602438947	c 358	19	2.4	632	10	BM335113	BM335113 MEST145-G
c 286	19	2.4	511	10	BF976030	BF976030 602244856	c 359	19	2.4	634	9	BB654539	BB654539 BB654539
c 287	19	2.4	511	10	BG341679	BG341679 602463402	c 360	19	2.4	636	9	AI055456	AI055456 coau0004A
c 288	19	2.4	513	10	N98756	N98756 zb85a06.s1	c 361	19	2.4	637	10	BM339332	BM339332 MEST240-A
c 289	19	2.4	514	9	AW185477	AW185477 se79a05.y	c 362	19	2.4	641	9	AW666939	AW666939 GA_Ea000
c 290	19	2.4	514	10	BI491296	BI491296 df06g02.w	c 363	19	2.4	642	12	BH123201	BH123201 RPC1-24-2
c 291	19	2.4	514	10	BM333138	BM333138 MEST184-E	c 364	19	2.4	646	9	BB391961	BB391961 BB391961
c 292	19	2.4	515	9	BB701448	BB701448	c 365	19	2.4	648	9	BB609370	BB609370 BB609370
c 293	19	2.4	516	10	BM421645	BM421645 V017G01 O	c 366	19	2.4	652	10	BM269225	BM269225 MEST406-F
c 294	19	2.4	516	10	BF551324	UI-R-C0-1	c 367	19	2.4	653	10	BM269225	BM269225 MEST406-F
c 295	19	2.4	519	10	BG811700	daf35e07.	c 368	19	2.4	654	12	BH124979	BH124979 RPC1-24-3
c 296	19	2.4	523	10	BI879984	BI879984 fm68c03.x	c 369	19	2.4	654	12	AQ874250	AQ874250 UI05D6.mt
c 297	19	2.4	524	12	A2992161	A2992161 2M0276007	c 370	19	2.4	658	10	BE907319	BE907319 601500028
c 298	19	2.4	525	12	AZ648929	AZ648929 1M0518H02	c 371	19	2.4	660	10	BJ011463	BJ011463 BJO11463
c 299	19	2.4	526	10	BG323192	BG323192 602421402	c 372	19	2.4	664	10	BG097566	BG097566 ES9462085
c 300	19	2.4	527	10	BI813671	BI813671 M002D02 O	c 373	19	2.4	666	9	BB622597	BB622597 BB622597
c 301	19	2.4	527	10	BM333306	BM333306 MEST188-H	c 374	19	2.4	668	12	AQ548213	AQ548213 RPC1-11-4
c 302	19	2.4	531	10	BM334919	BM334919 MEST130-A	c 375	19	2.4	669	12	AZ574127	AZ574127 323PvB07
c 303	19	2.4	531	10	BF515156	BF515156 UI-H-BW1-	c 376	19	2.4	672	10	BF786886	BF786886 602108815
c 304	19	2.4	532	9	AI752264	AI752264 cn14f05.y	c 377	19	2.4	672	12	AZ880966	AZ880966 RPC1-23-2
c 305	19	2.4	532	12	TA314F09Q	TA314F09Q	c 378	19	2.4	676	10	BG452759	BG452759 NF081B04L
c 306	19	2.4	536	12	AZ128698	AZ128698 OSJNB009	c 379	19	2.4	677	10	B1110864	B1110864 602895521
c 307	19	2.4	538	10	BM012778	BM012778 603637493	c 380	19	2.4	677	12	AG114870	AG114870 Pan trogl
c 308	19	2.4	538	10	BE519235	BE519235 946094G11	c 381	19	2.4	679	9	AV233538	AV233538 AV233538
c 309	19	2.4	541	10	BM419892	BM419892 R020C06 O	c 382	19	2.4	679	9	BB621608	BB621608 BB621608

383	19	2.4	680	9	BB496196	BB496196	456	18	2.3	146	9	AW102423	AW102423
c 384	19	2.4	681	10	BM268293	MEST379-C	c 457	18	2.3	153	10	BF413865	BF413865
c 385	19	2.4	681	12	BH546884	BOCHR46TR	c 458	18	2.3	156	10	BG796706	BG796706
c 386	19	2.4	682	10	BE457796	us97d06-x	c 459	18	2.3	158	12	AQ202583	AQ202583
c 387	19	2.4	683	10	BM336552	MEST195-E	c 460	18	2.3	159	9	AU165237	AU165237
c 388	19	2.4	686	10	BM074801	MEST297-G	c 461	18	2.3	164	10	BF632472	BF632472
c 389	19	2.4	687	10	BM268287	MEST379-B	c 462	18	2.3	176	12	AQ900972	AQ900972
c 390	19	2.4	687	10	BM334548	MEST138-E	c 463	18	2.3	180	9	BE007892	BE007892
c 391	19	2.4	689	10	B1335227	602998123	c 464	18	2.3	185	9	BB321552	BB321552
c 392	19	2.4	694	12	AZ703477	RPC1-23-2	c 465	18	2.3	186	9	AW588233	AW588233
c 393	19	2.4	696	12	B67072	CJT-HSP-201	c 466	18	2.3	189	9	AA014745	AA014745
c 394	19	2.4	697	12	AZ477093	1M0296H23	c 467	18	2.3	190	10	H64397	H64397
c 395	19	2.4	699	10	BM337919	MEST219-B	c 468	18	2.3	193	9	AV026747	AV026747
c 396	19	2.4	722	12	BH033898	RPC1-24-2	c 469	18	2.3	200	9	AI320464	AI320464
c 397	19	2.4	723	10	BM335663	MEST164-G	c 470	18	2.3	209	9	AV421167	AV421167
c 398	19	2.4	724	10	BM339178	MEST237-G	c 471	18	2.3	215	9	AV224372	AV224372
c 399	19	2.4	725	10	BM348599	MEST293-F	c 472	18	2.3	215	9	AA293052	AA293052
c 400	19	2.4	727	10	BM332094	MEST152-C	c 473	18	2.3	215	10	D23188	D23188
c 401	19	2.4	733	10	EG088921	H3159D02	c 474	18	2.3	216	9	AI311051	AI311051
c 402	19	2.4	738	10	EG206536	RST25987	c 475	18	2.3	216	9	AI311061	AI311061
c 403	19	2.4	745	10	BM049709	603624216	c 476	18	2.3	216	9	AI335306	AI335306
c 404	19	2.4	747	10	BM333846	MEST127-H	c 477	18	2.3	216	9	AV107276	AV107276
c 405	19	2.4	754	10	BE255574	601111779	c 478	18	2.3	219	9	BB074037	BB074037
c 406	19	2.4	768	10	BG077936	H3020F12-	c 479	18	2.3	220	9	AV163043	AV163043
c 407	19	2.4	776	12	AZ133005	OSJNB010	c 480	18	2.3	222	10	BM187060	BM187060
c 408	19	2.4	785	9	AI303541	u180q11-x	c 481	18	2.3	223	10	BM473610	BM473610
c 409	19	2.4	789	10	BE368703	601222667	c 482	18	2.3	227	9	AI320188	AI320188
c 410	19	2.4	794	10	BE961893	601648321	c 483	18	2.3	229	10	BF780651	BF780651
c 411	19	2.4	798	12	BH352429	CH230-33D	c 484	18	2.3	230	10	AI319780	AI319780
c 412	19	2.4	800	10	BE783117	601470716	c 485	18	2.3	234	9	AA676080	AA676080
c 413	19	2.4	805	10	BG398667	602440274	c 486	18	2.3	235	9	AE061495	AE061495
c 414	19	2.4	828	10	BG545588	602572843	c 487	18	2.3	235	12	AQ881790	AQ881790
c 415	19	2.4	830	10	BG755416	602713974	c 488	18	2.3	238	9	AV425036	AV425036
c 416	19	2.4	830	10	BE568447	601342532	c 489	18	2.3	238	12	BH509559	BH509559
c 417	19	2.4	847	10	EG180096	602329670	c 490	18	2.3	239	9	AV417837	AV417837
c 418	19	2.4	848	12	CNS035A0	AL228489 Tetraodon	c 491	18	2.3	239	9	AV236848	AV236848
c 419	19	2.4	849	10	BG250193	602362309	c 492	18	2.3	240	10	H64445	H64445
c 420	19	2.4	853	10	BF342443	602013948	c 493	18	2.3	243	9	AV232664	AV232664
c 421	19	2.4	858	10	BF123654	601760582	c 494	18	2.3	250	9	AV412809	AV412809
c 422	19	2.4	860	12	AQ273862	nbxb0031D	c 495	18	2.3	254	9	AW100656	AW100656
c 423	19	2.4	872	10	BI771964	BI771964 603058994	c 496	18	2.3	254	9	BB280343	BB280343
c 424	19	2.4	882	12	AZ207783	SP_0134_A	c 497	18	2.3	257	9	BB325964	BB325964
c 425	19	2.4	896	10	BF698958	602126547	c 498	18	2.3	257	10	BI278065	BI278065
c 426	19	2.4	898	9	BE052093	·BG244865 602358307	c 499	18	2.3	260	9	AI059224	AI059224
c 427	19	2.4	908	10	BE244865	·BG244865 602358307	c 500	18	2.3	260	9	AJ281544	AJ281544
c 428	19	2.4	930	12	CNS01S3C	AL164721 Tetraodon	c 501	18	2.3	260	9	BE059056	BE059056
c 429	19	2.4	931	10	BF163443	BF163443 601771612	c 502	18	2.3	265	9	AW529484	AW529484
c 430	19	2.4	956	12	CNS0272U	AL185331 Tetraodon	c 503	18	2.3	265	12	BH192911	BH192911
c 431	19	2.4	958	12	CNS06PBI	AL409139 T3 end of	c 504	18	2.3	266	10	BE953648	BE953648
c 432	19	2.4	976	12	CNS035YU	AL229359 Tetraodon	c 505	18	2.3	273	9	AA819089	AA819089
c 433	19	2.4	985	12	CNS037SC	AL231717 Tetraodon	c 506	18	2.3	273	9	AV417272	AV417272
c 434	19	2.4	988	10	BE907598	BE907598 601497593	c 507	18	2.3	274	10	BF371887	BF371887
c 435	19	2.4	989	9	A1964363	A1964363 EST269477	c 508	18	2.3	278	9	BB524171	BB524171
c 436	19	2.4	1010	10	BF203342	BF203342 601865946	c 509	18	2.3	278	12	AQ102939	AQ102939
c 437	19	2.4	1035	12	AG127413	AG127413 Pan trogl	c 510	18	2.3	282	9	AW493262	AW493262
c 438	19	2.4	1094	10	BF306609	BF306609 601888859	c 511	18	2.3	282	9	BF407424	BF407424
c 439	19	2.4	1146	10	EG113345	EG113345 602284950	c 512	18	2.3	284	9	BB079171	BB079171
c 440	19	2.4	1164	10	BE618333	BE618333 601462726	c 513	18	2.3	286	12	AZ4559363	AZ4559363
c 441	19	2.4	1292	10	BM415039	BM415039 6020095 M	c 514	18	2.3	288	9	AV555681	AV555681
c 442	19	2.4	1353	10	BE903463	BE903463 601677729	c 515	18	2.3	288	10	BI290964	BI290964
c 443	19	2.4	1403	12	AG060395	AG060395 Pan trogl	c 516	18	2.3	290	9	BB337259	BB337259
c 444	19	2.4	1685	10	BE733622	BE733622 601565994	c 517	18	2.3	290	9	BB373746	BB373746
c 445	19	2.4	1763	10	BF165204	BF165204 60178071	c 518	18	2.3	291	9	AV160063	AV160063
c 446	19	2.4	1827	10	BI489151	BI489151 603021259	c 519	18	2.3	291	9	AV427352	AV427352
c 447	19	2.4	1953	10	BF026022	BF026022 601670339	c 520	18	2.3	291	9	BB485672	BB485672
c 448	19	2.4	2160	10	EG024074	EG024074 602303119	c 521	18	2.3	291	10	BE476222	BE476222
c 449	18	2.3	118	9	AW491438	AW491438 UI-M-BH3-	c 522	18	2.3	292	10	BF610726	BF610726
c 450	18	2.3	122	9	AV415847	AV415847 AV415847	c 523	18	2.3	293	9	AI837074	AI837074
c 451	18	2.3	123	9	AV427993	AV427993 AV427993	c 524	18	2.3	293	9	AV412833	AV412833
c 452	18	2.3	125	10	BI671551	BI671551 fs49g05-y	c 525	18	2.3	293	9	BB555850	BB555850
c 453	18	2.3	132	10	BF469705	BF469705 UI-M-BH3-	c 526	18	2.3	296	9	AV118678	AV118678
c 454	18	2.3	140	10	BI714151	BI714151 1633e12.x	c 527	18	2.3	300	12	BH618245	BH618245
c 455	18	2.3	144	9	AW159612	AW159612 zb04g05.x	c 528	18	2.3	301	12	AQ095016	AQ095016

675 18 2.3 433 9 BE172782 MR0-HT055 748 18 2.3 495 9 AW500337
c 676 18 2.3 435 9 AW523833 UI-R-BJ0P c 749 18 2.3 496 10 BG632704
c 677 18 2.3 435 9 BE061545 MRO-BT024 750 18 2.3 496 10 BG632704
678 18 2.3 437 9 AW456686 UI-M-BH3- 751 18 2.3 497 12 AZ438612
679 18 2.3 440 9 AV406890 AV406890 752 18 2.3 497 9 AL119550
680 18 2.3 441 9 AV739205 AV739205 753 18 2.3 497 12 AZ719287
c 681 18 2.3 441 10 BG942075 ax20911.x 754 18 2.3 502 9 AI483446
682 18 2.3 442 9 AU182988 AU182988 755 18 2.3 502 10 BM383708
683 18 2.3 444 9 AV429135 AV429135 756 18 2.3 504 9 AI485030
c 684 18 2.3 447 9 AT072445 UI-R-C2-n 757 18 2.3 504 10 BG514253
c 685 18 2.3 447 9 BE061565 MRO-BT024 758 18 2.3 505 10 BI315894
c 686 18 2.3 448 9 AW938394 PM4-DT005 759 18 2.3 506 9 AI138041
c 687 18 2.3 448 9 BE061570 MRO-BT024 760 18 2.3 508 10 BE646204
c 688 18 2.3 448 9 BE061599 MRO-BT024 761 18 2.3 508 12 AQ167872
c 689 18 2.3 451 9 AT043907 UI-R-C0-J 762 18 2.3 510 10 BI298738
c 690 18 2.3 451 9 AW045648 UI-M-BH1- 763 18 2.3 511 9 AA437284
c 691 18 2.3 452 9 BE061594 MRO-BT024 764 18 2.3 511 10 H23095
c 692 18 2.3 452 9 BE099693 UI-R-BJ1- 765 18 2.3 511 12 AQ183788
c 693 18 2.3 452 10 BE448487 ut-90e05.y 766 18 2.3 512 10 BG685618
c 694 18 2.3 454 9 AV423526 AV423526 767 18 2.3 512 10 BG685618
c 695 18 2.3 454 9 AV763938 MRO-BT024 768 18 2.3 513 10 BE464656
c 696 18 2.3 454 9 BE061496 MRO-BT024 769 18 2.3 514 12 AZ223556
c 697 18 2.3 454 10 T47182 BI809999 1001A11.O 770 18 2.3 515 12 BE408313
c 698 18 2.3 454 10 T47182 BI809999 1001A11.O 771 18 2.3 515 12 AZ223556
c 699 18 2.3 455 9 AV1622350 AV423492 772 18 2.3 516 12 AQ392326
c 700 18 2.3 455 9 AV423492 AV423492 773 18 2.3 517 10 C96939
c 701 18 2.3 455 9 AW323244 uo57h06.y 774 18 2.3 517 12 AQ731916
c 702 18 2.3 456 9 AI314818 775 18 2.3 519 10 BF410145
c 703 18 2.3 456 9 AU181310 776 18 2.3 521 9 BE116156
c 704 18 2.3 456 9 BE146010 MRO-HT020 777 18 2.3 521 10 H08295
c 705 18 2.3 456 10 BF405710 778 18 2.3 522 9 AL598904
c 706 18 2.3 457 10 BF419597 UI-R-CA0- 779 18 2.3 525 9 AI733033
c 707 18 2.3 457 10 BF510908 UI-H-B14- 780 18 2.3 525 10 BM080277
c 708 18 2.3 458 9 AW036862 614019G07 781 18 2.3 525 12 AZ063489
c 709 18 2.3 459 9 AI397761 NCC5A477 782 18 2.3 525 12 AZ261475
c 710 18 2.3 459 9 BE146194 MRO-HT020 783 18 2.3 527 10 R60415
c 711 18 2.3 461 10 BI300463 784 18 2.3 528 9 AL380082
c 712 18 2.3 462 9 AI089922 qa16c08.x 785 18 2.3 528 10 BE628786
c 713 18 2.3 462 9 AV422721 AV422721 786 18 2.3 530 10 BE628786
c 714 18 2.3 462 9 AV423381 AV423381 787 18 2.3 531 10 H15326
c 715 18 2.3 463 10 BI810919 K011C02.O 788 18 2.3 531 12 AZ504695
c 716 18 2.3 463 12 AO83076 RPC111-56 789 18 2.3 533 10 BM127376
c 717 18 2.3 464 10 BF461828 UI-M-CG0P 790 18 2.3 533 10 H24417
c 718 18 2.3 465 10 BG399981 602442035 791 18 2.3 533 12 H24417
c 719 18 2.3 466 9 AT121737 fc31910.x 792 18 2.3 533 12 H24417
c 720 18 2.3 467 9 AW500274 793 18 2.3 533 12 H24417
c 721 18 2.3 467 10 BI809789 G002G03.O 794 18 2.3 533 12 H24417
c 722 18 2.3 468 10 R17192 Yq11d05.r1 795 18 2.3 533 12 H24417
c 723 18 2.3 470 10 BM419464 R014B04.O 796 18 2.3 533 12 H24417
c 724 18 2.3 471 9 BE146012 MRO-HT020 797 18 2.3 533 12 H24417
c 725 18 2.3 472 9 AV422843 798 18 2.3 533 12 H24417
c 726 18 2.3 474 9 BE146072 799 18 2.3 533 12 H24417
c 727 18 2.3 475 12 AZ429457 800 18 2.3 533 12 H24417
c 728 18 2.3 475 12 AZ657376 1M0533G04 801 18 2.3 533 12 H24417
c 729 18 2.3 478 9 AU095018 AU095018 802 18 2.3 533 12 H24417
c 730 18 2.3 478 10 BI118684 EST070.D1 803 18 2.3 533 12 H24417
c 731 18 2.3 478 12 AQ334072 HS_5010.A 804 18 2.3 533 12 H24417
c 732 18 2.3 480 10 H27819 Y162c04.r1 805 18 2.3 533 12 H24417
c 733 18 2.3 481 12 AQ280653 CITBI-E1- 806 18 2.3 533 12 H24417
c 734 18 2.3 482 9 BE760911 BB760911 807 18 2.3 533 12 H24417
c 735 18 2.3 482 10 BF393956 UI-R-CA0- 808 18 2.3 533 12 H24417
c 736 18 2.3 483 10 BF393956 UI-R-CA0- 809 18 2.3 533 12 H24417
c 737 18 2.3 483 10 BF386286 UI-R-CA1- 810 18 2.3 533 12 H24417
c 738 18 2.3 484 10 BG893169 daa89c04.x 811 18 2.3 533 12 H24417
c 739 18 2.3 484 10 BI841661 812 18 2.3 533 12 H24417
c 740 18 2.3 486 9 AA415106 813 18 2.3 533 12 H24417
c 741 18 2.3 488 9 AA899956 814 18 2.3 533 12 H24417
c 742 18 2.3 488 9 AA899956 815 18 2.3 533 12 H24417
c 743 18 2.3 488 10 R45475 Yq43f03.s1 816 18 2.3 533 12 H24417
c 744 18 2.3 491 12 AQ803076 HS_3161.B 817 18 2.3 533 12 H24417
c 745 18 2.3 494 9 AW282022 fj61b03.x 818 18 2.3 533 12 H24417
c 746 18 2.3 494 10 BI798252 H101H01.E 819 18 2.3 533 12 H24417
c 747 18 2.3 494 12 AQ346530 RPC111-10 820 18 2.3 533 12 H24417

AW500337 UI-HF-BNO
BG632704 GH12617.3
AZ438612 1M0228J14
AL119550 DKE2P761B
AZ719287 RPC1-24-1
A1483446 EST249267
BI383708 UI-R-DM1-
A1485030 EST243310
BG514253 G94904.x
BI315894 saf64902.
AI138041 UI-R-C2p-
BE646204 7683a06.x
BI167872 HS_2224.B
BI298738 UI-R-CV2-
AA437284 zv62b11.s
H23095 ym51d10.r1
AQ183788 HS_3199.B
BG685618 602637596
BE464656 hs86b02.x
AZ223556 RPC1-23-7
BE408313 601302313
AZ366375 1M0115005
AQ392326 CITBI-E1-
C96939 C96939 Rice
AQ731916 HS_5548.A
BF410145 UI-R-CA1-
BE116156 UI-R-BS1-
H08295 Y192b02.r1
AL598904 DKF2P313C
AI733033 oj28d01.x
BM080277 MEST106-D
AZ063489 RPC1-23-4
AZ261475 RPC1-23-1
R60415 yh04h05.r1
AL380082 MtBB50C09
BE628786 uu08a06.x
H15326 ym28d07.r1
AZ504695 1M0345D02
BM127376 1698a08.x
H24417 ym53f12.r1
B07921 pbAC 3.5 re
AO458440 HS_5207.A
BG181416 RST260.A
AL082297 Arad1d005
BM420236 0004G05.O
AQ784867 HS_3180.A
AZ271211 RPC1-23-1
BJ176566 BJ176566
BM379543 MEST507-A
AI398117 NCSM2R11T
AW128371 fe38f07.x
AI103454 EST212743
AI485939 EST244260
AW438183 707066E03
AO519123 HS_5150.B
AZ282429 RPC1-23-1
BE100176 UI-R-BJ1-
BI887123 ZF637-1-0
AQ965179 LERIA79TR
BI315597 CH230-196
AU145607 AU145607
AQ342749 RPC111-11
BE810064 218865.WA
BJ174299 BJ174299
BH197407 TC3-70D18
BM126792 IF03G06.x
AA778444 zf39h07.s
BM267876 MEST373-D
AL433780 Fugu rubr
H00142 Y16904.r1
BI798471 H108F07.E
BH503175 BOGGW70TR
AQ543840 RPC1-11-3

821	18	2.3	567	12	AQ543842	AQ543842	RPCI-11-3	894	18	2.3	665	10	BG622179	BG622179	602646817
822	18	2.3	569	9	AV855572	AV855572	AV855572	895	18	2.3	666	10	BE418909	BE418909	SCL083.H1
823	18	2.3	569	10	BG016317	BG016317	de59a11.x	896	18	2.3	666	12	BH340270	BH340270	CH230-891
824	18	2.3	569	10	B1796999	B1796999	H061B03.E	897	18	2.3	668	12	AG149209	AG149209	Pan trogl
825	18	2.3	569	10	AZ028973	AZ028973	RPCI-23-2	898	18	2.3	669	10	BJ001663	BJ001663	BJ001663
826	18	2.3	570	10	B1800002	B1800002	H145D11.E	899	18	2.3	669	10	BF650565	BF650565	NF098E04E
827	18	2.3	571	10	BM080235	BM080235	MEST105-H	900	18	2.3	670	9	AW687946	AW687946	NF0001C10S
828	18	2.3	571	10	BG456835	BG456835	NF098B10P	901	18	2.3	670	12	AG156264	AG156264	Pan trogl
829	18	2.3	572	10	B1810205	B1810205	J001E12	902	18	2.3	671	9	AW423087	AW423087	f168b12.Y
830	18	2.3	575	10	BG842783	BG842783	MEST10-40	903	18	2.3	671	12	AZ454449	AZ454449	IM0256123
831	18	2.3	575	10	BE435579	BE435579	EST406837	904	18	2.3	672	12	AZ358669	AZ358669	IM0101N08
832	18	2.3	576	12	AZ412911	AZ412911	IM0186109	905	18	2.3	673	9	AW691765	AW691765	NF049A04S
833	18	2.3	580	9	BE100641	BE100641	UI-R-BJ1-	906	18	2.3	673	10	BM412586	BM412586	EST586913
834	18	2.3	582	9	AW596699	AW596699	s115f05.Y	907	18	2.3	673	12	AZ572004	AZ572004	301PV808
835	18	2.3	583	10	B1804470	B1804470	H137G03.E	908	18	2.3	673	12	AG102439	AG102439	Pan trogl
836	18	2.3	584	10	BF255421	BF255421	HVSMF000	909	18	2.3	674	12	AG165186	AG165186	Pan trogl
837	18	2.3	584	12	BH452043	BH452043	B0HTA311F	910	18	2.3	676	9	AW918231	AW918231	EST349535
838	18	2.3	585	10	BF430890	BF430890	OG05G04T3	911	18	2.3	677	12	AG164122	AG164122	Pan trogl
839	18	2.3	589	10	B1806324	B1806324	S061E03.S	912	18	2.3	677	12	AQ643153	AQ643153	RPCI-11-3
840	18	2.3	589	10	BE948897	BE948897	UI-M-BH3-	913	18	2.3	683	12	AZ507099	AZ507099	IM0348M06
841	18	2.3	589	12	AQ586120	AQ586120	RPCI-11-4	914	18	2.3	685	10	BE917612	BE917612	930-f1a.S
842	18	2.3	590	9	AA210913	AA210913	tr90h07.Y	915	18	2.3	686	9	AU094199	AU094199	EST244979
843	18	2.3	591	12	AZ996520	AZ996520	2M0282B18	916	18	2.3	687	9	AI486658	AI486658	EST244979
844	18	2.3	591	12	AG071036	AG071036	Pan trogl	917	18	2.3	687	12	AQ537592	AQ537592	RPCI-11-3
845	18	2.3	592	10	B1521785	B1521785	603081767	918	18	2.3	688	12	AQ528621	AQ528621	RPCI-11-3
846	18	2.3	592	12	AQ930455	AQ930455	RPCI-23-2	919	18	2.3	692	12	BH323333	BH323333	CH230-132
847	18	2.3	594	9	AW291163	AW291163	UI-H-B12-	920	18	2.3	698	12	CNS04NOS	AL298981	Tetraodon
848	18	2.3	594	12	AQ969692	AQ969692	LERJP37TR	921	18	2.3	700	9	BE214327	BE214327	HV_CEB000
849	18	2.3	598	10	BG737810	BG737810	fo96409.Y	922	18	2.3	700	12	BH107913	BH107913	RPCI-24-2
850	18	2.3	601	9	AL635782	AL635782	AL635782	923	18	2.3	702	10	B1936679	B1936679	603089610
851	18	2.3	601	12	AQ551292	AQ551292	RPCI-11-4	924	18	2.3	705	12	BH477115	BH477115	BOGUP82TR
852	18	2.3	606	10	B1934590	B1934590	EST554479	925	18	2.3	707	10	BF480434	BF480434	L0-2368T13
853	18	2.3	608	10	BM385193	BM385193	UI-R-CN1-	926	18	2.3	707	12	AZ096850	AZ096850	RPCI-23-1
854	18	2.3	610	10	B1290550	B1290550	UI-R-DK0-	927	18	2.3	709	10	BG328939	BG328939	602428085
855	18	2.3	611	12	AQ635720	AQ635720	RPCI-11-4	928	18	2.3	710	12	AZ708896	AZ708896	RPCI-24-8
856	18	2.3	612	9	AI484060	AI484060	EST249931	929	18	2.3	715	12	AZ601349	AZ601349	IM0419N06
857	18	2.3	612	9	AW187450	AW187450	BNLGH1543	930	18	2.3	719	12	AZ723160	AZ723160	RPCI-24-9
858	18	2.3	612	10	B1417264	B1417264	LJNEST49D	931	18	2.3	720	12	AG001882	AG001882	Homo sapi
859	18	2.3	614	12	AQ555820	AQ555820	CITBI-E1-	932	18	2.3	726	12	AZ750682	AZ750682	RPCI-24-1
860	18	2.3	615	12	AQ965178	AQ965178	LERIA79TF	933	18	2.3	732	12	BH443907	BH443907	BOHRG25TF
861	18	2.3	616	12	AQ323307	AQ323307	RPCI11-10	934	18	2.3	733	10	BE410865	BE410865	601301196
862	18	2.3	618	10	B1965044	B1965044	id36b03.X	935	18	2.3	737	10	BG470332	BG470332	602704423
863	18	2.3	618	10	C91418	C91418	C91418.Dict	936	18	2.3	737	10	B1875882	B1875882	f164c06.Y
864	18	2.3	624	10	B1922170	B1922170	EST542073	937	18	2.3	742	12	AQ631229	AQ631229	RPCI-11-4
865	18	2.3	624	12	AZ367842	AZ367842	IM0117H09	938	18	2.3	745	10	BE390658	BE390658	601286829
866	18	2.3	625	9	AA660669	AA660669	00558.Mtr	939	18	2.3	748	12	AG181379	AG181379	Pan trogl
867	18	2.3	625	12	CNS01NIN	AL152308	Anopheles	940	18	2.3	751	12	AG186647	AG186647	Homo sapi
868	18	2.3	626	9	BB045047	BB045047	BB045047	941	18	2.3	757	10	BE513521	BE513521	601315028
869	18	2.3	626	12	AZ698024	AZ698024	RPCI-23-2	942	18	2.3	757	12	BH427300	BH427300	BOGOV51TR
870	18	2.3	628	10	BG189234	BG189234	RST8278.A	943	18	2.3	762	10	BG563260	BG563260	602582186
871	18	2.3	629	12	AZ262176	AZ262176	RPCI-23-1	944	18	2.3	765	12	AZ743739	AZ743739	RST124-7
872	18	2.3	631	12	AZ602680	AZ602680	IM0421K04	945	18	2.3	768	10	BG181921	BG181921	RST779.AC
873	18	2.3	632	10	BF977548	BF977548	602147327	946	18	2.3	769	10	BF232523	BF232523	de30f03.X
874	18	2.3	635	9	AV727200	AV727200	AV727200	947	18	2.3	779	9	BE113316	BE113316	UI-R-BJ1-
875	18	2.3	636	12	BH424853	BH424853	BOGOV49TR	948	18	2.3	780	9	AU130041	AU130041	AU130041
876	18	2.3	637	9	AL654707	AL654707	AL654707	949	18	2.3	781	12	BH497004	BH497004	BOGW232TR
877	18	2.3	639	9	AW255637	AW255637	ML696.pcp	950	18	2.3	789	12	BH273104	BH273104	CH230-83H
878	18	2.3	640	12	AZ364907	AZ364907	IM0111C22	951	18	2.3	790	10	BM412898	BM412898	EST587225
879	18	2.3	643	12	AZ218848	AZ218848	Sheared.D	952	18	2.3	794	10	BG567560	BG567560	602586313
880	18	2.3	645	10	BF029524	BF029524	601765926	953	18	2.3	794	12	CNS0225T	AL17770	Tetraodon
881	18	2.3	647	9	BE109607	BE109607	UI-R-BJ1-	954	18	2.3	798	9	AI668047	AI668047	TENG0925
882	18	2.3	649	12	AQ644424	AQ644424	RPCI93-EC	955	18	2.3	799	12	BH530996	BH530996	BOHLL88TR
883	18	2.3	650	9	BB638796	BB638796	BB638796	956	18	2.3	801	9	AI958297	AI958297	fd02h08.Y
884	18	2.3	650	10	BM419102	BM419102	R009D07.O	957	18	2.3	804	12	CNS01W1E	AL159835	Tetraodon
885	18	2.3	651	12	B13843	B13843	A-344C7.TP	958	18	2.3	807	10	BG843476	BG843476	1024002H1
886	18	2.3	654	9	AI484898	AI484898	EST243161	959	18	2.3	807	12	CNS04123	AL291684	Tetraodon
887	18	2.3	657	12	AZ091548	AZ091548	RPCI-23-2	960	18	2.3	809	10	B1737635	B1737635	603358544
888	18	2.3	659	9	BB478239	BB478239	BB478239	961	18	2.3	814	12	B20024	B20024	F4J12-T7.IG
889	18	2.3	660	12	AZ575044	AZ575044	335PVF09	962	18	2.3	818	9	AV645878	AV645878	AV645878
890	18	2.3	664	9	AW232538	AW232538	fj21e10.X	963	18	2.3	821	10	BG596211	BG596211	602659338
891	18	2.3	664	10	BG748799	BG748799	602706135	964	18	2.3	826	12	BH466068	BH466068	BOGHCL1TR
892	18	2.3	664	12	AG044826	AG044826	Pan trogl	965	18	2.3	827	9	AU142922	AU142922	AU142922
893	18	2.3	665	10	BG296106	BG296106	602393787	966	18	2.3	837	12	CNS02CNS	AL191629	Tetraodon

```
967 18 2.3 839 10 BF140226 BF140226 601788116
c 968 18 2.3 840 10 BE966747 BE966747 601661246
c 969 18 2.3 841 10 BI254450 BI254450 602976791
c 970 18 2.3 845 10 BF238816 BF238816 601904421
c 971 18 2.3 852 10 BG423027 BG423027 602450226
c 972 18 2.3 862 12 CNS01NOS AL152565 Anopheles
c 973 18 2.3 877 10 BF786559 BF786559 602108506
c 974 18 2.3 877 10 BE488164 BE488164 517700 Am
c 975 18 2.3 877 10 BF696784 BF696784 602125316
c 976 18 2.3 883 12 CNS008AW AL051442 Drosophila
c 977 18 2.3 884 12 CNS028DC AL185817 Tetraodon
c 978 18 2.3 888 12 CNS01NRA AL152577 Anopheles
c 979 18 2.3 891 12 CNS04WNR AL310572 Tetraodon
c 980 18 2.3 900 10 BF70911D AL343711 T3 end of
c 981 18 2.3 902 12 CNS051NV BI770911 603059723
c 982 18 2.3 905 12 BH135545 AL317092 Tetraodon
c 983 18 2.3 914 10 BE966270 BH135545 ENTOD19TF
c 984 18 2.3 936 12 CNS01NKS BE966270 601660510
c 985 18 2.3 936 12 CNS01NKS AL152349 Anopheles
c 986 18 2.3 941 10 BF121746 AL261421 Tetraodon
c 987 18 2.3 952 9 BE039984 BF121746 601759262
c 988 18 2.3 958 10 BE559666 BE039984 OC104H11
c 989 18 2.3 960 10 BE736499 BE559666 601347393
c 990 18 2.3 963 12 CNS0301P BF736499 601306440
c 991 18 2.3 968 12 CNS07137 AL221686 Tetraodon
c 992 18 2.3 972 10 BF526115 AL244409 T3 end of
c 993 18 2.3 974 10 BF311887 BF526115 602071012
c 994 18 2.3 980 12 CNS05NE6 BF311887 601897721
c 995 18 2.3 996 12 CNS04NRR AL345255 Tetraodon
c 996 18 2.3 999 12 BG492592 AL299070 Tetraodon
c 997 18 2.3 1000 12 AG086210 BG492592 602536556
c 998 18 2.3 1001 12 CNS03BHL AG086210 Pan trogl
c 999 18 2.3 1005 12 CNS01TOA AL236514 Tetraodon
c1000 18 2.3 1005 12 CNS01TOA AL166771 Tetraodon
```

ALIGNMENTS

```
RESULT 1
BE875341 785 bp mRNA linear EST 20-OCT-2000
LOCUS 601489051F1 NIH_MGC_69 Homo sapiens cDNA clone IMAGE:3891444 5',
DEFINITION mRNA sequence.
ACCESSION BE875341
VERSION BE875341.1 GI:10324117
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheraia; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: DCTD/DBP/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9676 row: e column: 13
High quality sequence stop: 725.
Location/Qualifiers
1. .785
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3891444"
/clone_lib="NIH_MGC_69"
/tissue_type="large cell carcinoma, undifferentiated"
```

```
/lab_host="DH10B (phage-resistant)"
/notes="organ: lung; Vector: pcMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.1 kb. Library constructed by Life
Technologies."
BASE COUNT 245 a 173 c 234 g 133 t
ORIGIN
Query Match 66.7%; Score 521; DB 10; Length 785;
Best Local Similarity 100.0%; Pred. No. 5.4e-178;
Matches 521; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 40 CTGGATTCTTCTCGCTACTGAGACGCGGGGTAGGTCCACAGGAGATCCCACTGGG 99
Db 1 CTGGATTCTTCTCGCTACTGAGACGCGGGGTAGGTCCACAGGAGATCCCACTGGG 60
QY 100 AGTTGAAGTGTGAGTGAGAGTAAGAGGAACAGCAGAGCTTCCCGAGGGTGTGTGTC 159
Db 61 AGTTGAAGTGTGAGTGAGAGTAAGAGGAACAGCAGAGCTTCCCGAGGGTGTGTGTC 120
QY 160 GTGACTCAGATGAGAAAGCCCTCGAAGTCGTCGCTCTCATCGCGTGCACGCCCAT 219
Db 121 GTGACTCAGATGAGAAAGCCCTCGAAGTCGTCGCTCTCATCGCGTGCACGCCCAT 180
QY 220 GGACCTTCTTCTCGTCTCACGGCCATAACTAGGGAGGAGGCGCGAGGAGTGAGGG 279
Db 181 GGACCTTCTTCTCGTCTCACGGCCATAACTAGGGAGGAGGCGCGAGGAGTGAGGG 240
QY 280 GCTCAGGCGAAGCTGGGGTGTGTTGGGGGTATCCGAGTCCCAAGACCTTGGAAACCC 339
Db 241 GCTCAGGCGAAGCTGGGGTGTGTTGGGGGTATCCGAGTCCCAAGACCTTGGAAACCC 300
QY 340 GACAGAAGATTCTGACATCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 399
Db 301 GACAGAAGATTCTGACATCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
QY 400 CAAACA*AGAACCCACAGCCAGTCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 459
Db 361 CAAACA*AGAACCCACAGCCAGTCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
QY 460 AACCGAGCTGAAAGTCTGGGATCTACACCTGGGAGGAGGAGGAGGAGGAGGAGGAGG 519
Db 421 AACCGAGCTGAAAGTCTGGGATCTACACCTGGGAGGAGGAGGAGGAGGAGGAGGAGG 480
QY 520 CAGCTGAGATCCCGAGTGGCGGAGCATGGAAGTGTGTCGAA 560
Db 481 CAGCTGAGATCCCGAGTGGCGGAGCATGGAAGTGTGTCGAA 521
```

```
RESULT 2
BE881023 489 bp mRNA linear EST 20-OCT-2000
LOCUS 601490780F1 NIH_MGC_69 Homo sapiens cDNA clone IMAGE:3893227 5',
DEFINITION mRNA sequence.
ACCESSION BE881023
VERSION BE881023.1 GI:10329799
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheraia; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: DCTD/DBP/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
```


KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
DEFINITION Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 487)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10774 row: p column: 19
High quality sequence stop: 483.
Location/Qualifiers
1. .487
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4839426"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.2 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 166 a 113 c 123 g 85 t
ORIGIN
Query Match 49.9%; Score 390; DB 10; Length 487;
Best Local Similarity 100.0%; Pred. No. 1.1e-130;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 392 GACACACAAACACAGAACCCACAGCCAGTCCCAGGAGCCCAAGTAAATGGAGAGCCCA 451
Db 81 GACACACAAACACAGAACCCACAGCCAGTCCCAGGAGCCCAAGTAAATGGAGAGCCCA 140
QY 452 AAAGAAGAACACAGAGCTGAAGTCGGATCCTACACCTGGCGACGACACAGAACAGA 511
Db 141 AAAGAAGAACACAGAGCTGAAGTCGGATCCTACACCTGGCGACGACAGAACAGA 200
QY 512 TCAGGATACAGCTGAGATCCAGTCGGCGACATGGAAGGTGATCTGCAAGAGCTGCATCA 571
Db 201 TCAGGATACAGCTGAGATCCAGTCGGCGACATGGAAGGTGATCTGCAAGAGCTGCATCA 260
QY 572 GTCAACACCGGGGATAAATCTGGATTGGTTCCGGCGTCAAGGTGAAGATAATACCTA 631
Db 261 GTCAACACCGGGGATAAATCTGGATTGGTTCCGGCGTCAAGGTGAAGATAATACCTA 320
QY 632 AAGAGGAACACTGTAATAATGCCAAGACAGGTGAAGAGCAACACAAAGTTTAAATGAAGA 691
Db 321 AAGAGGAACACTGTAATAATGCCAAGACAGGTGAAGAGCAACACAAAGTTTAAATGAAGA 380
QY 692 CAAGCTGAACACAGCAAGCTGTTTTATATTAGATATTTTGACTTAAACTATCTCAATAA 751
Db 381 CAAGCTGAACACAGCAAGCTGTTTTATATTAGATATTTTGACTTAAACTATCTCAATAA 440
QY 752 AGTTTTGCAGCTTTTCACCAAAAAA 781
Db 441 AGTTTTGCAGCTTTTCACCAAAAAA 470

RESULT 5
BE876234
LOCUS
DEFINITION BE876234 880 bp mRNA linear EST 20-OCT-2000
601486758F1 NIH_MGC_69 Homo sapiens cDNA clone IMAGE:3889095 5',
mRNA sequence.
ACCESSION BE876234
VERSION BE876234.1 GI:10325114
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
DEFINITION Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 880)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/BTP/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9670 row: c column: 16
High quality sequence stop: 493.
Location/Qualifiers
1. .880
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3889095"
/clone_lib="NIH_MGC_69"
/tissue_type="large cell carcinoma, undifferentiated"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: PCW-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.1 kb. Library constructed by Life
Technologies."
BASE COUNT 382 a 192 c 209 g 97 t
ORIGIN
Query Match 49.9%; Score 390; DB 10; Length 880;
Best Local Similarity 100.0%; Pred. No. 8.1e-131;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 392 GACACACAAACACAGAACCCACAGCCAGTCCCAGGAGCCCAAGTAAATGGAGAGCCCA 451
Db 64 GACACACAAACACAGAACCCACAGCCAGTCCCAGGAGCCCAAGTAAATGGAGAGCCCA 123
QY 452 AAAGAAGAACACAGAGCTGAAGTCGGATCCTACACCTGGCGACGACAGAACAGA 511
Db 124 AAAGAAGAACACAGAGCTGAAGTCGGATCCTACACCTGGCGACGACAGAACAGA 183
QY 512 TCAGGATACAGCTGAGATCCAGTCGGCGACATGGAAGGTGATCTGCAAGAGCTGCATCA 571
Db 184 TCAGGATACAGCTGAGATCCAGTCGGCGACATGGAAGGTGATCTGCAAGAGCTGCATCA 243
QY 572 GTCAACACCGGGGATAAATCTGGATTGGTTCCGGCGTCAAGGTGAAGATAATACCTA 631
Db 244 GTCAACACCGGGGATAAATCTGGATTGGTTCCGGCGTCAAGGTGAAGATAATACCTA 303
QY 632 AAGAGGAACACTGTAATAATGCCAAGACAGGTGAAGAGCAACACCAAGTTTAAATGAAGA 691
Db 304 AAGAGGAACACTGTAATAATGCCAAGACAGGTGAAGAGCAACCAAGTTTAAATGAAGA 363
QY 692 CAAGCTGAACACAGCAAGCTGTTTTATATTAGATATTTTGACTTAAACTATCTCAATAA 751
Db 364 CAAGCTGAACACAGCAAGCTGTTTTATATTAGATATTTTGACTTAAACTATCTCAATAA 423
QY 752 AGTTTTGCAGCTTTTCACCAAAAAA 781
Db 424 AGTTTTGCAGCTTTTCACCAAAAAA 453

```

RESULT 6
AA948244/c
LOCUS
DEFINITION
  AA948244      428 bp      mRNA      linear      EST 04-MAY-1998
  oq33901.s1 NCI_CGAP_GC4 Homo sapiens cDNA clone IMAGE:1588176 3',
  mRNA sequence.
ACCESSION
  AA948244
VERSION
  AA948244.1 GI:3109497
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 428)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
  Emmert-Buck, M.D., Ph.D.
  cDNA Library Preparation: M. Bento Soares, Ph.D.
  cDNA Library Arrayed by: Greg Lennon, Ph.D.
  DNA Sequencing by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/fbrp/image/image.html
  Seq primer: -40ml3 fwd. ET from Amersham.
  Location/Qualifiers
    1..428
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:1588176"
    /clone_lib="NCI_CGAP_GC4"
    /tissue_type="pooled germ cell tumors"
    /lab_host="DH10B"
    /note="vector: pT7T3D-Pac (Pharmacia) with a modified
    polylinker; 1st strand cDNA was prepared from 3 pooled
    germ cell tumors, and was then primed with a Not I -
    oligo(dT) primer. Double-stranded cDNA was ligated to Eco
    RI adaptors (Pharmacia), digested with Not I and cloned
    into the Not I and Eco RI sites of the modified pT7T3
    vector. Library is normalized. Library was constructed by
    Bento Soares and M. Fatima Bonaldo."
    BASE COUNT      80 a 104 c 99 g 145 t
    ORIGIN
      Query Match      48.5%; Score 379; DB 9; Length 428;
      Best Local Similarity 100.0%; Pred. No. 1.1e-126;
      Mismatches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

  QY 392 GACACACAAACACAGAACCCACAGCCAGTCCAGGAGCCCGAGTATGGAGAGCCCCA 451
  Db 379 GACACACAAACACAGAACCCACAGCCAGTCCAGGAGCCCGAGTATGGAGAGCCCCA 320

  QY 452 AAAGAGAACACAGCAGCTGAAAGTCGGGATCCTACACCTGGCGAGCAGACAGAGA 511
  Db 319 AAAGAGAACACAGCAGCTGAAAGTCGGGATCCTACACCTGGCGAGCAGACAGAGA 260

  QY 512 TCAGGATACAGCTGAGATCCCAAGTCGGGCACATGGAAGTGATCTGCAAGAGCTGCATCA 571
  Db 259 TCAGGATACAGCTGAGATCCCAAGTCGGGCACATGGAAGTGATCTGCAAGAGCTGCATCA 200

  QY 572 GTCAACACCGGGGATAAATCTGGATTTGGGTTCCGGGCTCAAGGTGAAGATAAATACCTA 631
  Db 199 GTCAACACCGGGGATAAATCTGGATTTGGGTTCCGGGCTCAAGGTGAAGATAAATACCTA 140

  QY 632 AAGAGAACACTGTAAATGCCAAGAGCAGGTGAAGAGCAACACCAAGTTTAAATGAAGA 691
  Db 139 AAGAGAACACTGTAAATGCCAAGAGCAGGTGAAGAGCAACACCAAGTTTAAATGAAGA 80

```

```

QY 692 CAAGCTGAACAAACGCAAGCTGTTTTATATTAGATATTTGACTTAACTATCTCAATAA 751
Db 79 CAAGCTGAACAAACGCAAGCTGTTTTATATTAGATATTTGACTTAACTATCTCAATAA 20

QY 752 AGTTTTCAGCTTTTCACCA 770
Db 19 AGTTTTCAGCTTTTCACCA 1

RESULT 7
AA620697/c
LOCUS
DEFINITION
  af89401.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1049185
  3', mRNA sequence.
ACCESSION
  AA620697
VERSION
  AA620697.1 GI:2524636
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 414)
  Hillier L., Allen M., Bowles L., Dubuque T., Geisel G., Jost S.,
  Krizman D., Kucaba T., Lacy M., Le N., Lennon G., Marra M., Martin
  J., Moore B., Schellenberg K., Steptoe M., Tan F., Theising B.,
  White Y., Wyllie T., Waterston R. and Wilson R.
  WashU-NCI human EST Project
  Unpublished (1997)
  Contact: Wilson RK
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: est@watson.wustl.edu
  This clone is available royalty-free through LLNL; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  Seq primer: -40ml3 fwd. ET from Amersham
  High quality sequence stop: 375.
  Location/Qualifiers
    1..414
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:1049185"
    /clone_lib="Soares_testis_NHT"
    /sex="male"
    /lab_host="DH10B"
    /note="vector: pT7T3D-Pac (Pharmacia) with a modified
    polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
    was prepared from mRNA obtained from Clontech Laboratories
    , Inc., and primed with a Not I - oligo(dT) primer [5',
    TGTTACCAATCTGAAGTGGCGCGCGCCCAATTTTTTTTTTTT 3'].
    Double-stranded cDNA was ligated to Eco RI adaptors
    (Pharmacia), digested with Not I and cloned into the Not I
    and Eco RI sites of the modified pT7T3 vector. Library
    went through one round of normalization to Col5, and was
    constructed by Bento Soares and M. Fatima Bonaldo."
    BASE COUNT      77 a 100 c 94 g 143 t
    ORIGIN
      Query Match      48.4%; Score 378; DB 9; Length 414;
      Best Local Similarity 100.0%; Pred. No. 2.5e-126;
      Mismatches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

  QY 392 GACACACAAACACAGAACCCACAGCCAGTCCAGGAGCCCGAGTATGGAGAGCCCCA 451
  Db 378 GACACACAAACACAGAACCCACAGCCAGTCCAGGAGCCCGAGTATGGAGAGCCCCA 319

  QY 452 AAAGAGAACACAGCAGCTGAAAGTCGGGATCCTACACCTGGCGAGCAGACAGAGA 511
  Db 318 AAAGAGAACACAGCAGCTGAAAGTCGGGATCCTACACCTGGCGAGCAGACAGAGA 259

  QY 512 TCAGGATACAGCTGAGATCCCAAGTCGGGCACATGGAAGTGATCTGCAAGAGCTGCATCA 571
  Db 139 TCAGGATACAGCTGAGATCCCAAGTCGGGCACATGGAAGTGATCTGCAAGAGCTGCATCA 571

```

```

Db 258 TCAGGATACAGCTGAGATCCAGTCCGCGCATGGAAGTGATCTGCAAGAGCTGCATCA 199
QY 572 GTCAACACCGGGGATAAATCTCGATTGGTTCCGCGCTCAAGGTGAAGATAATACCTA 631
Db 198 GTCAACAACGCGGGATAAATCTCGATTGGTTCCGCGCTCAAGGTGAAGATAATACCTA 139
QY 632 AAGAGGAACACTGTAAATGCCAAGACAGCTGAAGAGCAACACAAAGTTTAAATGAAGA 691
Db 138 AAGAGGAACACTGTAAATGCCAAGACAGCTGAAGAGCAACACAAAGTTTAAATGAAGA 79
QY 692 CAAGCTGAACACGCAAGCTGGTTTATATTAGATATTGACTTAAACTATCTCAATAA 751
Db 78 CAAGCTGAACACGCAAGCTGGTTTATATTAGATATTGACTTAAACTATCTCAATAA 19
QY 752 AGTTTTCAGCTTTCACC 769
Db 18 AGTTTTCAGCTTTCACC 1

RESULT 8
AI954712/c
LOCUS      wq32c10.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2472978 3',
DEFINITION      mRNA sequence.
ACCESSION      AI954712.1 GI:5747022
VERSION        AI954712.1
KEYWORDS       EST.
SOURCE         human.
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 438)
AUTHORS        NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE          National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
               Tumor Gene Index
JOURNAL         Unpublished (1997)
COMMENT        Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-re@mail.nih.gov
               Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
               R. Emmert-Buck, M.D., Ph.D.
               cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
               Bonaldo, Ph.D.
               DNA Library Arrayed by: Greg Lennon, Ph.D.
               DNA Sequencing by: Washington University Genome Sequencing Center
               Clone distribution: NCI-CGAP clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               www.bio.llnl.gov/dbbrp/image/image.html
               Insert Length: 536 Std Error: 0.00
               Seq primer: -40UP from Gibco.
               Location/Qualifiers
                 1..438
                   /organism="Homo sapiens"
                   /db_xref="taxon:9606"
                   /clone="IMAGE:2472978"
                   /clone_lib="NCI_CGAP_GC6"
                   /tissue_type="pooled germ cell tumors"
                   /lab_host="DH10B"
                   /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                   polylinker. Site.1: Not I; Site.2: Eco RI; Plasmid DNA
                   from the normalized library NCI_CGAP_GC4 was prepared, and
                   ss circles were made in vitro. Following HAP purification,
                   this DNA was used as tracer in a subtractive hybridization
                   reaction. The driver was PCR-amplified cDNAs from a pool
                   of 5,000 clones made from the same library (cloneIDs
                   1257096-1258631, 1469064-1470983, and 1475592-1476743).
                   Subtraction by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      79 a 104 c 97 g 158 t
ORIGIN
Query Match      47.0%; Score 367; DB 9; Length 438;
Best Local Similarity 100.0%; Pred. No. 2.3e-122;
Matches 367; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 392 GACACACAAAACACAGAACACACAGCCAGTCCCGAGAGCCAGTAATGGAGAGCCCCA 451
Db 391 GACACACAAAACACAGAACACACAGCCAGTCCCGAGAGCCAGTAATGGAGAGCCCCA 332
QY 452 AAAAGAAGAACACAGCAGCTGAAAGTCGGATCCTACACCTGGGCAGCAGACAGAAGA 511
Db 331 AAAAGAAGAACACAGCAGCTGAAAGTCGGATCCTACACCTGGGCAGCAGACAGAAGA 272
QY 512 TCAGGATACAGCTGAGATCCAGTCCGCGACATGGAAGTGATCTGCAAGAGCTGCATCA 571
Db 271 TCAGGATACAGCTGAGATCCAGTCCGCGACATGGAAGTGATCTGCAAGAGCTGCATCA 212
QY 572 GTCAACACCGGGGATAAATCTGGATTTCGGCTCCCGCTCAAGGTGAAGATAATACCTA 631
Db 211 GTCAACACCGGGGATAAATCTGGATTTCGGCTCCCGCTCAAGGTGAAGATAATACCTA 152
QY 632 AAGAGGAACACTGTAAATGCCAAGACAGCTGAAGAGCAACACAAAGTTTAAATGAAGA 691
Db 151 AAGAGGAACACTGTAAATGCCAAGACAGCTGAAGAGCAACACAAAGTTTAAATGAAGA 92
QY 692 CAAGCTGAACACGCAAGCTGGTTTATATTAGATATTGACTTAAACTATCTCAATAA 751
Db 91 CAAGCTGAACACGCAAGCTGGTTTATATTAGATATTGACTTAAACTATCTCAATAA 32
QY 752 AGTTTTCAGCTTTCACC 758
Db 31 AGTTTTCAGCTTTCACC 25

RESULT 9
BE620196
LOCUS      601483262F1 NIH_MGC_69 Homo sapiens cDNA clone IMAGE:388585 5',
DEFINITION      mRNA sequence.
ACCESSION      BE620196
VERSION        BE620196.1 GI:9891134
KEYWORDS       EST.
SOURCE         human.
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 1140)
AUTHORS        NIH-MGC http://mgc.nci.nih.gov/
TITLE          National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL         Unpublished (1999)
COMMENT        Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-re@mail.nih.gov
               Tissue Procurement: DCTD/DRP/Gazdar
               cDNA Library Preparation: Life Technologies, Inc.
               DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: Incyte Genomics, Inc.
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov
               Plate: LLAM9661 row: 1 column: 16
               High quality sequence stop: 440.
               Location/Qualifiers
                 1..1140
                   /organism="Homo sapiens"
                   /db_xref="taxon:9606"
                   /clone="IMAGE:388585"
                   /clone_lib="NIH_MGC_69"
                   /tissue_type="large cell carcinoma, undifferentiated"
                   /lab_host="DH10B (phage-resistant)"
                   /note="Organ: lung; Vector: PCMV-SPORT6; Site.1: NotI;
                   Site.2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                   Average insert size 1.1 kb. Library constructed by Life
                   Technologies."
BASE COUNT      540 a 242 c 266 g 90 t
ORIGIN
Query Match      46.1%; Score 360; DB 10; Length 1140;
Best Local Similarity 100.0%; Pred. No. 4.6e-120;

```

Matches 360; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	408	GAACACACAGCCAGTCCAGGAGCCAGTAAATGGAGAGCCCAAAAGAAAGAACACGCA	467						
Db	1	GAACACACAGCCAGTCCAGGAGCCAGTAAATGGAGAGCCCAAAAGAAAGAACACGCA	60						
QY	468	GCTGAAAGTCGGGATCTACACCTCGGCGAGCAGACAGAAAGATCAGGATACAGCTGAG	527						
Db	61	GCTGAAAGTCGGGATCTACACCTCGGCGAGCAGACAGAAAGATCAGGATACAGCTGAG	120						
QY	528	ATCCAGTCGCGGACATGGAAGTGATCTGCAAGAGCTGCATCAGTCAAAACCGGGGAT	587						
Db	121	ATCCAGTCGCGGACATGGAAGTGATCTGCAAGAGCTGCATCAGTCAAAACCGGGGAT	180						
QY	588	AAATCTGGATTTGGGTTTCGGCGCTCAAGGTGAAGTAATACTCAAGAGAAACACTGTAA	647						
Db	181	AAATCTGGATTTGGGTTTCGGCGCTCAAGGTGAAGTAATACTCAAGAGAAACACTGTAA	240						
QY	648	AATGCCAGAGCAGGTGAAGAGCAACCAACAGTTTAAATGAAGCAACAGCTGAAACACGC	707						
Db	241	AATGCCAGAGCAGGTGAAGAGCAACCAACAGTTTAAATGAAGCAACAGCTGAAACACGC	300						
QY	708	AAGCTGGTTTTATATTAGATATTGACTTAAACTATCTCAATAAAGTTTTCAGCTTTTCA	767						
Db	301	AAGCTGGTTTTATATTAGATATTGACTTAAACTATCTCAATAAAGTTTTCAGCTTTTCA	360						
RESULT 10									
AT631681/c									
LOCUS									
DEFINITION									
wb08d02.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2305059 3', mRNA sequence.									
ACCESSION									
AT631681									
VERSION									
AT631681.1 GI:4683011									
KEYWORDS									
EST.									
SOURCE									
human.									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
AUTHORS									
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .									
TITLE									
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index									
JOURNAL									
COMMENT									
Unpublished (1997)									
Contact: Robert Strausberg, Ph.D.									
Email: cgaps-remail.nih.gov									
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.									
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.									
cDNA Library Arrayed by: Greg Lennon, Ph.D.									
DNA Sequencing by: Washington University Genome Sequencing Center									
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/brp/image/image.html									
Insert Length: 515 Std Error: 0.00									
Seq primer: -400P from Gibco.									
FEATURES									
source									
1. .421									
Location/Qualifiers									
/organism="Homo sapiens"									
/db_xref="taxon:9606"									
/clone="IMAGE:2305059"									
/clone_lib="NCI_CGAP_GC6"									
/tissue_type="pooled germ cell tumors"									
/lab_host="DH10B"									
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_GC4 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1257096-1258631, 1469064-1470983, and 1475592-1476743).									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
492									
BASE COUNT									
168 a 125 c 145 g 118 t									
ORIGIN									
Query Match 43.3%; Score 338; DB 9; Length 556;									
Best Local Similarity 100.0%; Pred. No. 5.6e-112;									
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY 433 CCAGTAAATGGAGAGCCCAAAAGAAAGAACACAGCAGCTGAAAGTCGGGATCTACACCTG									
49									


```

Db 94  CCAGTAATGAGAGCCCAAAAAAGAAAGAACACGACGCTGAAAGTCGGGATCCTACACCTG 153
      |||
Qy 493  GGCAGCAGACAGAGAAGATCAGGATACAGCTCAGATCCAGTGCAGCATGGAAGGTG 552
      |||
Db 154  GGCAGCAGACAGAGAAGATCAGGATACAGCTCAGATCCAGTGCAGCATGGAAGGTG 213
      |||
Qy 553  ATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTTGGGTTCCGGCGTC 612
      |||
Db 214  ATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTTGGGTTCCGGCGTC 273
      |||
Qy 613  AAGGTGAAGATAATACCTAAGAGGACACTGTAATGTCAGCAAGCAGTGAAGACAA 672
      |||
Db 274  AAGGTGAAGATAATACCTAAGAGGACACTGTAATGTCAGCAAGCAGTGAAGACAA 333
      |||
Qy 673  CCACAAGTTTAAATGAAGACAAGCTGAAACAACGCAAGCTGGTTTATATTAGATATTG 732
      |||
Db 334  CCACAAGTTTAAATGAAGACAAGCTGAAACAACGCAAGCTGGTTTATATTAGATATTG 393
      |||
Qy 733  ACTTAAACTATCTCAATAAAGTTTTCGAGCTTTCACCA 770
      |||
Db 394  ACTTAAACTATCTCAATAAAGTTTTCGAGCTTTCACCA 431
      |||

RESULT 12
AA648117/c  AA648117 423 bp mRNA linear EST 13-NOV-1997
LOCUS ns06407.r1 NCI_CGAP_Ew1 Homo sapiens cDNA clone IMAGE:1182829, mRNA
sequence.
ACCESSION AA648117
VERSION AA648117.1 GI:2574546
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (bases 1 to 423)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lee Helman, M.D., Michael R. Emmert-Buck, M.D.,
Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 599 Std Error: 0.00
Seq primer: -28ml3 rev1 Et from Amersham
High quality sequence stop: 408.
FEATURES
source
1. .423
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1182829"
/clone_lib="NCI_CGAP_Ew1"
/tissue_type="Ewing's sarcoma"
/lab_host="DH10B"
/notes="Vector: pAMP10; mRNA made from Ewing's sarcoma,
cDNA made by oligo-dT priming. Non-directionally cloned.
Size-selected on agarose gel, average insert size 600 bp.
Reference: Krizman et al. (1996) Cancer Research
56:5380-5383."
BASE COUNT 77 a 101 c 90 g 155 t
ORIGIN

Query Match 42.9%; Score 335; DB 9; Length 423;
Best Local Similarity 99.7%; Pred. No. 7.8e-111;
Matches 385; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

Qy 392  GACACACACAAACACAGAACCCACACAGCCAGTCCCAGGAGCCAGTAATGGAGAGCCCCA 451
      |||
Db 415  GACACACACAAACACAGAACCCACACAGCCAGTCCCAGGAGCCAGTAATGGAGAGCCCCA 356
      |||
Qy 452  AAAGAAGAANCAGCAGCTGAAGTCGGATCCTACACCTGGCAGCAGACAGAAGA 511
      |||
Db 355  AACAGAAGAANCAGCAGCTGAAGTCGGATCCTACACCTGGCAGCAGACAGAAGA 295
      |||
Qy 512  TCAGGATACAGCTGAGATCCAGTCGCGACATGGAAGGTGATCTCCAAGAGCTGCATCA 571
      |||
Db 295  TCAGGATACAGCTGAGATCCAGTCGCGACATGGAAGGTGATCTCCAAGAGCTGCATCA 236
      |||
Qy 572  GTCAAACACCCGGGATAAATCTGGATTTGGGTTCCGGCTCAAGGTGAAGATAAATACCTA 631
      |||
Db 235  GTCAAACACCCGGGATAAATCTGGATTTGGGTTCCGGCTCAAGGTGAAGATAAATACCTA 176
      |||
Qy 632  AAGAGGAACACTGTAAATGCCAGAAGCAGGTGAAGCAGCAACCAAGTTTAAATCAAGA 691
      |||
Db 175  AAGAGGAACACTGTAAATGCCAGAAGCAGGTGAAGCAGCAACCAAGTTTAAATCAAGA 116
      |||
Qy 692  CAAGCTGAAACAACGCAAGCTGGTTTATATTAGATATTGACTTAAACTATCTCAATAA 751
      |||
Db 115  CAAGCTGAAACAACGCAAGCTGGTTTATATTAGATATTGACTTAAACTATCTCAATAA 56
      |||
Qy 752  AGTTTTCGAGCTTTCACCAAAAAA 777
      |||
Db 55  AGTTTTCGAGCTTTCACCAAAAAA 30
      |||

RESULT 13
AA758618/c  AA758618 407 bp mRNA linear EST 23-JAN-1998
LOCUS ah58d07.s1 Soares_testis_NHT Homo sapiens cDNA clone 1309837 3',
sequence.
ACCESSION AA758618
VERSION AA758618.1 GI:2806481
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (bases 1 to 407)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40ml3 fwd. Et from Amersham
High quality sequence stop: 401.
FEATURES
source
1. .407
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="1309837"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker: Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, inc., and primed with a Not I - oligo(dT) primer [5',
TCTTACCAATCTGAAGTGGAGCCGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
```

and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 77 a 97 c 91 g 142 t
ORIGIN

Query Match 39.7%; Score 310; DB 9; Length 407;
Best Local Similarity 100.0%; Pred. No. 8.1e-102;
Matches 310; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 GACACACAAACACAGAACACACAGCCAGTCCAGGAGCCCACTAATGAGAGCCCCA 451

Db 378 GACACACAAACACAGAACACACAGCCAGTCCAGGAGCCCACTAATGAGAGCCCCA 319

QY 452 AAAAGAAAGAACCCAGCAGCTGAAAGTCGGGATCCTACACCTCGGCAGCAGACAGAGAAGA 511

Db 318 AAAAGAAAGAACCCAGCAGCTGAAAGTCGGGATCCTACACCTCGGCAGCAGACAGAGAAGA 259

QY 512 TCAGGATACAGCTGAGATCCAGTCGCGGACATGGAAGGTGATCTGCAAGAGCTGCATCA 571

Db 258 TCAGGATACAGCTGAGATCCAGTCGCGGACATGGAAGGTGATCTGCAAGAGCTGCATCA 199

QY 572 GTCACACACCGGGGATAATCTGGATTTCGGCTTCGGCGCTCAAGGTGAAGATAATACCTA 631

Db 198 GTCACACACCGGGGATAATCTGGATTTCGGCTTCGGCGCTCAAGGTGAAGATAATACCTA 139

QY 632 AAGAGGAACACTGTAAAATGCCAGAAGCAGGTGAAGAGCAACCACTTTAAATGAAGA 691

Db 138 AAGAGGAACACTGTAAAATGCCAGAAGCAGGTGAAGAGCAACCACTTTAAATGAAGA 79

QY 692 CAAGCTGAAA 701

Db 78 CAAGCTGAAA 69

RESULT 14
BG217244
LOCUS BG217244 485 bp mRNA linear EST 21-APR-2001
DEFINITION RST36946 Athersys RAGE Library Homo sapiens cdNA, mRNA sequence.
ACCESSION BG217244
VERSION BG217244.1 GI:13743265
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 485)
Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R., Cain,S., Leventhal,C., Thornton,M., Ramchandran,R., Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,k., Offenbacher,J., Danzig,J. and Ducar,M.
Creation of genome-wide protein expression libraries using random activation of gene expression

TITLE
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE 2127151
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atersys.com
High quality sequence stop: 481.

FEATURES
source 1. 485
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/cell_line="HT1080"
/note="See 'Creation of Genome-wide protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation

method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

BASE COUNT 172 a 105 c 115 g 93 t
ORIGIN

Query Match 36.9%; Score 288; DB 10; Length 485;
Best Local Similarity 99.5%; Pred. No. 6.3e-94;
Matches 388; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 392 GACACACAAACACAGAACACACAGCCAGTCCAGGAGCCCACTAATGAGAGCCCCA 451

Db 95 GACACACAAACACAGAACACACAGCCAGTCCAGGAGCCCACTAATGAGAGCCCCA 154

QY 452 AAAAGAAAGAACCCAGCAGCTGAAAGTCGGGATCCTACACCTCGGCAGCAGACAGAGAAGA 511

Db 155 AAAAGAAAGAACCCAGCAGCTGAAAGTCGGGATCCTACACCTCGGCAGCAGACAGAGAAGA 214

QY 512 TCAAGATACAGCTGAGATCCAGTCGCGGACATGGAAGGTGATCTGCAAGAGCTGCATCA 571

Db 215 TCAAGATACAGCTGAGATCCAGTCGCGGACATGGAAGGTGATCTGCAAGAGCTGCATCA 274

QY 572 GTCACACACCGGGGATAATCTGGATTTCGGCTTCGGCGCTCAAGGTGAAGATAATACCTA 631

Db 275 GTCACACACCGGGGATAATCTGGATTTCGGCTTCGGCGCTCAAGGTGAAGATAATACCTA 334

QY 632 AAGAGGAACACTGTAAAATGCCAGAAGCAGGTGAAGAGCAACCACTTTAAATGAAGA 691

Db 335 AAGAGGAACACTGTAAAATGCCAGAAGCAGGTGAAGAGCAACCACTTTAAATGAAGA 394

QY 692 CAAGCTGAAAACACGCAAGCTGGCTTTATATAGATATTTCGACTTAACTATCTCAATAA 751

Db 395 CAAGCTGAAAACACGCAAGCTGGCTTTATATAGATATTTCGACTTAACTATCTCAATAA 454

QY 752 AGTTTTCAGCTTTTCACCAAAAAA 781

Db 455 AGTTTTCAGCTTTTCACCAAAAAA 484

RESULT 15
AV682444
LOCUS AV682444 665 bp mRNA linear EST 16-JAN-2002
DEFINITION AV682444 GKB Homo sapiens cdNA clone GKBACC10 5', mRNA sequence.
ACCESSION AV682444
VERSION AV682444.1 GI:10284307
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 665)
Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,O., Cai,T., Zhang,X., Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W., Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X., Hu,G., Gu,J., Chen,Z. and Han,Z.

TITLE
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE 21625106
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers
source 1. .665
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GKBACC10"
/clone_lib="GKB"

```
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/notes=Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT      212 a 152 c 152 g 149 t
ORIGIN
Query Match      36.0%; Score 281; DB 9; Length 665;
Best Local Similarity 100.0%; Pred. No. 1.8e-91;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 487 CACCTGGGCAGACAGAGAAGATCAGGATACAGCTGAGATCCACTGCGGCACATGG 546
DB 1 CACCTGGGCAGACAGAGAAGATCAGGATACAGCTGAGATCCACTGCGGCACATGG 60
QY 547 AAGGTGATCTCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTCGATTGGGTTC 606
DB 61 AAGGTGATCTCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTCGATTGGGTTC 120
QY 607 GCGTCAAGTCAAGATATACCTTAAGAGGAACACTGTAATGCCAGAGCAGGTCAA 666
DB 121 GCGTCAAGTCAAGATATACCTTAAGAGGAACACTGTAATGCCAGAGCAGGTCAA 180
QY 667 GAGCAACCAAGTTTAAATGAAGCAAGCTGAAACACGCAAGCTGTTTATATTAGA 726
DB 181 GAGCAACCAAGTTTAAATGAAGCAAGCTGAAACACGCAAGCTGTTTATATTAGA 240
QY 727 TATTGACTTAACTATCTCAATAAGTTTTCACGCTTCA 767
DB 241 TATTGACTTAACTATCTCAATAAGTTTTCACGCTTCA 281
RESULT 16
BG216048 394 bp mRNA linear EST 21-APR-2001
LOCUS RST35862 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
DEFINITION BG216048
ACCESSION BG216048.1 GI:13742197
VERSION EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 394)
AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.,
Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith
,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
,J., Danzig,J. and Ducar,M.
TITLE Creation of genome-wide protein expression libraries using random
activation of gene expression
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE 21227151
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@athersys.com
High quality sequence stop: 394.
Location/Qualifiers
1. 394
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/cell_line="HT1080"
/notes="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."
```

```
BASE COUNT      149 a 85 c 92 g 68 t
ORIGIN
Query Match      35.1%; Score 274; DB 10; Length 394;
Best Local Similarity 100.0%; Pred. No. 7.7e-89;
Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 440 TGGAGAGCCCAAAAAGAACACGAGCTGAAAGTCGGGATCCTACACCTGGGCAGCA 499
DB 52 TGGAGAGCCCAAAAAGAACACGAGCTGAAAGTCGGGATCCTACACCTGGGCAGCA 111
QY 500 GACAGAAGAAGATCAGGATCAGATCCCGCAGCTGCGGCACATGGAAGTGATCTGCA 559
DB 112 GACAGAAGAAGATCAGGATCAGATCCCGCAGCTGCGGCACATGGAAGTGATCTGCA 171
QY 560 AGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTGGGTTCGGCGTCAAGGTGA 619
DB 172 AGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTGGGTTCGGCGTCAAGGTGA 231
QY 620 AGATAATACCTTAAGAGGAACACTGTAATGCCAGAGCAGGTGAAGAGCAACCAAG 679
DB 232 AGATAATACCTTAAGAGGAACACTGTAATGCCAGAGCAGGTGAAGAGCAACCAAG 291
QY 680 TTTAAATGAAGACAAGCTGAAACAACGCAAGCTG 713
DB 292 TTTAAATGAAGACAAGCTGAAACAACGCAAGCTG 325
RESULT 17
BG206865 388 bp mRNA linear EST 21-APR-2001
LOCUS RST26327 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
DEFINITION BG206865
ACCESSION BG206865.1 GI:13728552
VERSION EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 388)
AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.,
Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith
,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
,J., Danzig,J. and Ducar,M.
TITLE Creation of genome-wide protein expression libraries using random
activation of gene expression
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE 21227151
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@athersys.com
High quality sequence stop: 228.
Location/Qualifiers
1. 388
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/cell_line="HT1080"
/notes="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."
```

```
BASE COUNT      143 a 83 c 87 g 71 t
ORIGIN
Query Match      33.2%; Score 259; DB 10; Length 388;
Best Local Similarity 100.0%; Pred. No. 2e-83;
```

```
Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 473 AAGTCGGGATCTACACCTGGGCAGCAGACAGAAGATCAGGATACAGCTGAGATCCC 532
      |||||||
Db 78 AAGTCGGGATCTACACCTGGGCAGCAGACAGAAGATCAGGATACAGCTGAGATCCC 137
      |||||||
Qy 533 AGTCGCGACATGGAGGTGATCTCCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATC 592
      |||||||
Db 138 AGTCGCGCAGATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATC 197
      |||||||
Qy 593 TGGATTGGGTTCCGCGCTCAAGGTGAAGATAATACCTAAAGAGGAACACTGTAAATGTC 652
      |||||||
Db 198 TGGATTGGGTTCCGCGCTCAAGGTGAAGATAATACCTAAAGAGGAACACTGTAAATGTC 257
      |||||||
Qy 653 CAGAAGCAGGTGAAGAGCAACACCAAGTTTAATGAAGACACAGCTGAACACACGCAAGCT 712
      |||||||
Db 258 CAGAAGCAGGTGAAGAGCAACACCAAGTTTAATGAAGACACAGCTGAACACACGCAAGCT 317
      |||||||
Qy 713 GGTTTTATATAGATATTT 731
      |||||||
Db 318 GGTTTTATATAGATATTT 336

RESULT 18
AA528460
LOCUS
DEFINITION
  AA528460 392 bp mRNA linear EST 19-AUG-1997
  similar to contains Alu repetitive element.; mRNA sequence.
ACCESSION
  AA528460
VERSION
  AA528460.1 GI:2270529
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 392)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
JOURNAL
  Unpublished (1997)
COMMENT
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: Lee Helman, M.D., Michael R. Emmert-Buck, M.D.,
  Ph.D.
  CDNA Library Preparation: David B. Krizman, Ph.D.
  CDNA Library Arrayed by: Greg Lennon, Ph.D.
  DNA Sequencing by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/bbrp/image/image.html
  Insert Length: 472 Std Error: 0.00
  Seq primer: -40m13 fwd. ET from Amersham
  High quality sequence stop: 350.
  Location/Qualifiers
    1..392
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="IMAGE:910989"
      /clone_lib="NCI-CGAP_Ew1"
      /tissue_type="Ewing's sarcoma"
      /lab_host="DH10B"
      /note="Vector: pAMP10; mRNA made from Ewing's sarcoma,
      cDNA made by oligo-dT priming. Non-directionally cloned.
      Size-selected on agarose gel, average insert size 600 bp.
      Reference: Krizman et al. (1996) Cancer Research
      56:5380-5383."
BASE COUNT 137 a 65 c 87 g 103 t
ORIGIN
  Query Match 31.5%; Score 246; DB 9; Length 392;
  Best Local Similarity 100.0%; Pred. No. 9,5e-79;
  Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 533 AGTCGCGACATGGAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATC 592
      |||||||
Db 126 AGTCGCGCAGCATGGAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATC 185
      |||||||
Qy 593 TGGATTGGGTTCCGCGCTCAAGGTGAAGATAATACCTAAAGAGGAACACTGTAAATGTC 652
      |||||||
Db 186 TGGATTGGGTTCCGCGCTCAAGGTGAAGATAATACCTAAAGAGGAACACTGTAAATGTC 245
      |||||||
Qy 653 CAGAAGCAGGTGAAGAGCAACACCAAGTTTAATGAAGACACAGCTGAACACACGCAAGCT 712
      |||||||
Db 246 CAGAAGCAGGTGAAGAGCAACACCAAGTTTAATGAAGACACAGCTGAACACACGCAAGCT 305
      |||||||
Qy 713 GGTTTTATATAGATATTTGACTTAAACTATCTCAATAAAGTTTTTGCAGCTTTCACCAA 772
      |||||||
Db 306 GGTTTTATATAGATATTTGACTTAAACTATCTCAATAAAGTTTTTGCAGCTTTCACCAA 365
      |||||||
Qy 773 AAAAAA 778
      |||||||
Db 366 AAAAAA 371

RESULT 19
AA483082
LOCUS
DEFINITION
  AA483082 369 bp mRNA linear EST 14-AUG-1997
  similar to contains Alu repetitive element.; mRNA sequence.
ACCESSION
  AA483082
VERSION
  AA483082.1 GI:2211927
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 369)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
JOURNAL
  Unpublished (1997)
COMMENT
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: Lee Helman, M.D., Michael R. Emmert-Buck, M.D.,
  Ph.D.
  CDNA Library Preparation: David B. Krizman, Ph.D.
  CDNA Library Arrayed by: Greg Lennon, Ph.D.
  DNA Sequencing by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/bbrp/image/image.html
  Insert Length: 506 Std Error: 0.00
  Seq primer: -41m13 fwd. ET from Amersham
  High quality sequence stop: 355.
  Location/Qualifiers
    1..369
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="IMAGE:909733"
      /clone_lib="NCI-CGAP_Ew1"
      /tissue_type="Ewing's sarcoma"
      /lab_host="DH10B"
      /note="vector: pAMP10; mRNA made from Ewing's sarcoma,
      cDNA made by oligo-dT priming. Non-directionally cloned.
      Size-selected on agarose gel, average insert size 600 bp.
      Reference: Krizman et al. (1996) Cancer Research
      56:5380-5383."
BASE COUNT 123 a 67 c 80 g 99 t
ORIGIN
  Query Match 31.1%; Score 243; DB 9; Length 369;
  Best Local Similarity 100.0%; Pred. No. 1.2e-77;
  Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



```
Db 208 CAGAGCAGGTGAAGAGCAACCAACGCTTTAAATGAAGACAGCTGAACCAACGCAAGCT 267
QY 713 GCTTTTATTAGATTGACTTAACTATCTCAATAAAGTTTTGGCAGCTTTCA 767
|||||
Db 268 GCTTTTATTAGATTGACTTAACTATCTCAATAAAGTTTTGGCAGCTTTCA 322
|||||

RESULT 22
AW386187 230 bp mRNA linear EST 04-FEB-2000
LOCUS RCO-PT0006-271199-011-E04 PT0006 Homo sapiens cDNA, mRNA sequence.
DEFINITION AW386187
ACCESSION AW386187.1 GI:8890846
VERSION AW386187.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 230)
HCGP http://www.ludwig.org.br/ORESTES.
AUTHORS The FAPESP/LICR Human Cancer Genome Project
TITLE Unpublished (1999)
JOURNAL Laboratory of Cancer Genetics
COMMENT Contact: Simpson A.J.G.
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC0&t2=RC0-PT0006-
271199-011-E04&t3=1999-11-27&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 230.
FEATURES
Location/Qualifiers
1..230
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="PT0006"
/dev_stage="Adult"
/note="organ: pnet; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT 78 a 52 c 65 g 35 t
ORIGIN
Query Match 29.4%; Score 230; DB 9; Length 230;
Best Local Similarity 100.0%; Pred. No. 7.3e-73;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 424 CCAGAGGCCAGTATGAGAGCCCCCAAAAAGAGACACCACTGAAAGTCGGGATC 483
|||||
Db 1 CCCAGGAGGCCAGTATGAGAGCCCCCAAAAAGAGACACCACTGAAAGTCGGGATC 60
|||||
QY 484 CTACCTGGGCAGCAGACAGAGAGATCAGGATACAGCTCCAGTCCGCGACA 543
|||||
Db 61 CTACCTGGGCAGCAGACAGAGAGATCAGGATACAGCTCCAGTCCGCGACA 120
|||||
QY 544 TGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTGGGT 603
|||||
Db 121 TGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTGGGT 180
|||||
QY 604 TCCGGCGTCAAGGTGAAGATAATACCTAAAGAGGACACTGTAATGCC 653
|||||
Db 181 TCCGGCGTCAAGGTGAAGATAATACCTAAAGAGGACACTGTAATGCC 230
|||||
```

```
RESULT 23
AA642878 318 bp mRNA linear EST 27-OCT-1997
LOCUS nu05H04.s1 NCI_CGAP_Alvl Homo sapiens cDNA clone IMAGE:1207159,
DEFINITION mRNA sequence.
ACCESSION AA642878
VERSION AA642878.1 GI:2568096
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 318)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lee Helman, M.D., Michael R. Emmert-Buck, M.D.,
Ph.D.
CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 312.
FEATURES
Location/Qualifiers
1..318
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:1207159"
/clone_lib="NCI_CGAP_Alvl"
/tissue_type="alveolar rhabdomyosarcoma"
/lab_host="DH10B"
/note="vector: pAMP10; mRNA made from alveolar
rhabdomyosarcoma, cDNA made by oligo-dr priming.
Non-directionally cloned. Size-selected on agarose gel,
average insert size 600 bp. Reference: Krizman et al.
(1996) Cancer Research 56:5380-5383."
BASE COUNT 110 a 51 c 72 g 85 t
ORIGIN
Query Match 29.4%; Score 230; DB 9; Length 318;
Best Local Similarity 100.0%; Pred. No. 6.2e-73;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 533 AGTGGCCACATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAACACCGGGGATAAATC 592
|||||
Db 89 AGTGGCCACATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAACACCGGGGATAAATC 148
|||||
QY 593 TGGATTTGGGTTCCGCGCTCAAGGTGAAGATAATACCTAAAGAGCAACACCTGTAATATGC 652
|||||
Db 149 TGGATTTGGGTTCCGCGCTCAAGGTGAAGATAATACCTAAAGAGCAACACCTGTAATATGC 208
|||||
QY 653 CAGAAGCAGGTGAAGAGCAACCAACGTTTAAATGAAGACAGCTGAACCAACCGCAAGCT 712
|||||
Db 209 CAGAAGCAGGTGAAGAGCAACCAACGTTTAAATGAAGACAGCTGAACCAACCGCAAGCT 268
|||||
QY 713 GGTTTTATTAGATTGACTTAACTATCTCAATAAAGTTTTTGCAGC 762
|||||
Db 269 GGTTTTATTAGATTGACTTAACTATCTCAATAAAGTTTTTGCAGC 318
|||||

RESULT 24
BG191783 1248 bp mRNA linear EST 21-APR-2001
LOCUS RST10879 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
DEFINITION BG191783
ACCESSION BG191783.1 GI:13713470
KEYWORDS EST.
```

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Kucuba,T., Lacy,M., Le.N., Lennon,G., Marra,M., Martin,J., Moore,B.,
AUTHORS Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wyllie
1 (bases 1 to 1248)
Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.,
Lerner,L., Costanzo,D., McElligott,K., Boozer,S., Mays,K., Smith
,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo.K., Offenbacher
,J., Danzig,J. and Bucar,M.
TITLE Creation of genome-wide protein expression libraries using random
activation of gene-wide expression
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE 21227151
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave. Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atersys.com
High quality sequence stop: 301.
Location/Qualifiers
1. .1248
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/cell_line="Ht1080"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is Ht1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in Ht1080 under normal circumstances."
BASE COUNT 291 a 278 c 257 g 401 t 21 others
ORIGIN
Query Match 27.4%; Score 214; DB 10; Length 1248;
Best Local Similarity 99.6%; Pred. No. 1.8e-67;
Matches 264; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 473 AAGTCGGGATCCTACACCTGGGCAGCAGACAGAAGAAGATCAGGATACAGCTGAGATCCC 532
|||||
Db 129 AAGTCGGGATCCTACACCTGGGCAGCAGACAGAAGAAGATCAGGATACAGCTGAGATCCC 188
QY 533 AGTCGGCGACATGAGAGTGATCTGCAAGAGCTGCATCAGTCAACACACCGGGGATAAATC 592
|||||
Db 189 AGTCGGCGACATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAACACACCGGGGATAAATC 248
QY 593 TGGATTGGGTTCCGGCGCTCAAGGTGAAGATAATACCTAAAGAGGAACACTCTAAATGC 652
|||||
Db 249 TGGATTGGGTTCCGGCGCTCAAGGTGAAGATAATACCTAAAGAGGAACACTCTAAATGC 308
QY 653 CAGAAGCAGGTGAAGAGCAACCAAGTTAAATGAAGCAAGCTGAAACACACGCAAGCT 712
|||||
Db 309 CAGAAGCAGGTGAAGAGCAACCAAGTTAAATGAAGCAAGCTGAAACACACGCAAGCT 368
QY 713 GGTTTATATTAGATATTTGACTTA 737
|||||
Db 369 GGTTTATATTAGATATTTGACTTA 393
RESULT 25
AA412359
LOCUS 2199a04.r1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:730446
DEFINITION 5', mRNA sequence.
ACCESSION AA412359
VERSION AA412359.1 GI:2070929
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
AUTHORS Kucuba,T., Lacy,M., Le.N., Lennon,G., Marra,M., Martin,J., Moore,B.,
, Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wyllie
,T., Waterston,R. and Wilson,R.
TITLE WashU-Merck EST Project 1997
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 503 Std Error: 0.00
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 256.
Location/Qualifiers
1. .299
/organism="Homo sapiens"
/db_xref="GDB:5927024"
/db_xref="taxon:9606"
/clone="IMAGE:730446"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 86 a 83 c 88 g 42 t
ORIGIN
Query Match 26.2%; Score 205; DB 9; Length 299;
Best Local Similarity 100.0%; Pred. No. 6.5e-64;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 392 GACACACAAACACAGAACACACAGCAGTCCAGAGCCCGCCAGATGGAGAGCCCA 451
|||||
Db 95 GACACACAAACACAGAACACACAGCAGTCCAGAGCCCGCCAGATGGAGAGCCCA 154
QY 452 AAAAGAAGAACACAGCAGCTGAAAGTCGGGATCCTACACCTGGGCAGCAGACAGAAGA 511
|||||
Db 155 AAAAGAAGAACACAGCAGCTGAAAGTCGGGATCCTACACCTGGGCAGCAGACAGAAGA 214
QY 512 TCAGGATACAGTGAGATCCCAGTCCGCGACATGGAAGGTGATCTGCAAGAGCTGCATCA 571
|||||
Db 215 TCAGGATACAGTGAGATCCCAGTCCGCGACATGGAAGGTGATCTGCAAGAGCTGCATCA 274
QY 572 GTCAAACACCGGGGATAAATCTGGA 596
|||||
Db 275 GTCAAACACCGGGGATAAATCTGGA 299
RESULT 26
AA747475
LOCUS nx73C03.s1 NCL_CCAP_Ew1 Homo sapiens cDNA clone IMAGE:1267876, mRNA
DEFINITION sequence.
ACCESSION AA747475
VERSION AA747475.1 GI:2787433
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

BASE COUNT		77 a		63 c		47 g		102 t	
ORIGIN		/tissue_type="alveolar rhabdomyosarcoma" /lab_host="DH10B" /note="Vector: pAMP10; mRNA made from alveolar rhabdomyosarcoma, cDNA made by oligo-dT priming. Non-directionally cloned. Size-selected on agarose gel, average insert size 600 bp. Reference: Krizman et al. (1996) Cancer Research 56:5380-5383." (996) Cancer Research 56:5380-5383."							
QY	533	AGTGC	CGCACATG	GAGGTG	ATCTG	CAAGAGCT	GCATC	AGTCAACAC	CGGGGATAATC 592
Db	240	AGTGC	CGCACATG	GAGGTG	ATCTG	CAAGAGCT	GCATC	AGTCAACAC	CGGGGATAATC 181
QY	593	TGGATT	TGGTTCCG	CGCGTCA	AGGTGA	AGATAATAC	TAAAGAGG	AACTGT	AAAAATGC 652
Db	180	TGGATT	TGGTTCCG	CGCGTCA	AGGTGA	AGATAATAC	TAAAGAGG	AACTGT	AAAAATGC 121
QY	653	CAGA	ACGAGTGA	AGAGCA	ACCACA	AGTTTAA	TATGAAG	CAACAGCT	GAAACACGCAAGCT 712
Db	120	CAGA	ACGAGTGA	AGAGCA	ACCACA	AGTTTAA	TATGAAG	CAACAGCT	GAAACACGCAAGCT 61
QY	713	GGTTT	TATA 721						
Db	60	GGTTT	TATA 52						
RESULT 30 AA730288 LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapbs@rmail.nih.gov Tissue Procurement: Lee Helman, M.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: David B. Krizman, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CCGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/dbbrp/image/image.html Insert Length: 281 Std Error: 0.00 Seq primer: -40m13 fwd. ET from Amersham.									
FEATURES source 1. .202 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:1249152" /clone_lib="NCI_CGAP_Ew1" /tissue_type="Ewing's sarcoma" /lab_host="DH10B" /note="Vector: pAMP10; mRNA made from Ewing's sarcoma, cDNA made by oligo-dT priming. Non-directionally cloned. Size-selected on agarose gel, average insert size 600 bp. Reference: Krizman et al. (1996) Cancer Research 56:5380-5383."									

```
BASE COUNT      79 a      34 c      41 g      48 t
ORIGIN
Query Match      22.2%; Score 173; DB 9; Length 202;
Best Local Similarity 100.0%; Pred. No. 2.7e-52;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 607 GCGTCAGGTGAAGTAAATACCTAAAGAGGAACACTGTAAATGCCAGAGCAGGTGAA 666
      |||
DB 23 GCGTCAGGTGAAGTAAATACCTAAAGAGGAACACTGTAAATGCCAGAGCAGGTGAA 82
      |||
QY 667 GAGCAACACAAAGTTTAAATGAAGCAACAGCTGAAACCAACGCAAGCTGTTTATATTAGA 726
      |||
DB 83 GAGCAACCAAGTTTAAATGAAGCAACAGCTGAAACCAACGCAAGCTGTTTATATTAGA 142
      |||
QY 727 TATTGACTTAACTATCTCAATAAAGTTTTCAGCTTTCCACCAAAAAA 779
      |||
DB 143 TATTGACTTAACTATCTCAATAAAGTTTTCAGCTTTCCACCAAAAAA 195
      |||

RESULT 31
A1205793/c      291 bp      mRNA      linear      EST 29-NOV-1998
LOCUS      qg33b06.x1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1836947
DEFINITION      3', mRNA sequence.
ACCESSION      A1205793
VERSION      A1205793.1 GI:3764465
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 291)
AUTHORS      NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaide
, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 555 Std Error: 0.00
Seq primer: -40up from Gibco
High quality sequence stop: 248.
Location/Qualifiers
1..291
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1836947"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - Oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGGAGCGCGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaide."
BASE COUNT      60 a      68 c      52 g      111 t
ORIGIN
Query Match      22.2%; Score 173; DB 9; Length 291;
Best Local Similarity 99.3%; Pred. No. 2.2e-52;
Matches 273; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

BASE COUNT      79 a      34 c      41 g      48 t
ORIGIN
Query Match      20.9%; Score 163; DB 10; Length 487;
Best Local Similarity 99.2%; Pred. No. 6.9e-49;
Matches 263; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 392 GACACACAAACACACACACACAGCCAGTCCCGAGGCCAGTAAATGAGAGCCCCA 451
      |||
DB 94 GACACACAAACACACACACACAGCCAGTCCCGAGGCCAGTAAATGAGAGCCCCA 153
      |||
```

```
QY 493 GGCAGCAGACAGAAGAAGATCAGGATACAGCTGAGATCCAGTCCGCGACATGAAAGGTG 552
      |||
DB 290 GGCAGCAGACAGAAGAAGATCAGGATACAGCTGAGATCCAGTCCGCGACATGAAAGGTG 231
      |||
QY 553 ATCTGAAGAGCTGCATCAGTCAACACCGGGGATAAATCTGGATTGGTTCGGCGCTC 612
      |||
DB 230 ATCTGAAGAGCTGCATCAGTCAACACCGGGGATAAATCTGGATTGGTTCGGCGCTC 171
      |||
QY 613 AAGTGAAGATAATACCTAAAGAGGAACACTGTAAATGCCAGAGCAGGTGAAGAGCAA 672
      |||
DB 170 AAGTGAAGATAATACCTAAAGAGGAACACTGTAAATGCCAGAGCAGGTGAAGAGCAA 111
      |||
QY 673 CCACAAGTTTAAATGAAGCAACAGCTGAAACCAACGCAAGCTGTTTATATTAGATTTG 732
      |||
DB 110 CCCAAGTTTAAATGAAGCAACAGCTGAAACCAACGCAAGCTGTTTATATTAGATTTG 51
      |||
QY 733 ACTTAACTATCTCAATAAAGTTTTCAGCTTTCA 767
      |||
DB 50 ACTTAACTATCTCAATAAAGTTTTCAGCTTTCA 16
      |||

RESULT 32
BG211375      487 bp      mRNA      linear      EST 21-APR-2001
LOCUS      RST30937 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
DEFINITION      BG211375
ACCESSION      BG211375
VERSION      BG211375.1 GI:13733062
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 487)
AUTHORS      Hall, Upton, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,
Cain, S., Leventhal, C., Thornton, M., Ramachandran, R., Whittington, J.,
Lerner, L., Costanzo, D., McElligott, K., Booser, S., Mays, R., Smith
, E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K., Offenbacher
, J., Danzig, J. and Ducar, M.
TITLE      Creation of genome-wide protein expression libraries using random
activation of gene expression
JOURNAL      Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE      21227151
COMMENT      Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@athersys.com
High quality sequence stop: 343.
Location/Qualifiers
1..487
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/clone_line="HT1080"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression', the
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."
BASE COUNT      169 a      100 c      127 g      91 t
ORIGIN
Query Match      20.9%; Score 163; DB 10; Length 487;
Best Local Similarity 99.2%; Pred. No. 6.9e-49;
Matches 263; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 452 AAAAGACAACCCAGCTGAAAGTCCGGATCCTACACCTGGGCAGCAGACAGAAGA 511
Db 154 AAAAGAGAAGACAGCAGCTGAAAGTCCGGATCCTACACCTGGGCAGCAGACAGAAGA 213

QY 512 TCAGGATACAGCTGAGATCCCGAGTCGCGCAGCATGGAAGGTGATCTGCAAGAGCTGCATCA 571
Db 214 TCAGGATACAGCTGTTATCCAGTCGCGCAGCATGGAAGGTGATCTGCAAGAGCTGCATCA 273

QY 572 GTCAAACACCGGGGATAAATCTGGATTGGTTCCGGGTCGAAGGTGAAGATAATACCTA 631
Db 274 GTCAAACACCGGGGATAAATCTGGATTGGTTCCGGGTCGAAGGTGAAGATAATACCTA 333

QY 632 AAGAGGAACACTGTAATAATGCCAGA 656
Db 334 AAGAGGAACACTGTAATAATGCCAGA 358

RESULT 33
LOCUS BG205627 391 bp mRNA linear EST 21-APR-2001
DEFINITION RST25074 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG205627
VERSION BG205627.1 GI:13727330
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 391)
AUTHORS Harrington,J.,J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.,
Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith
,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
,J., Danzig,J. and Ducar,M.
TITLE Creation of genome-wide protein expression libraries using random
activation of gene expression
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE 21227151
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scai@atersys.com
High quality sequence stop: 391.
FEATURES
source
1..391
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/cell_line="HT1080"
Note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."
BASE COUNT 144 a 84 c 90 g 73 t
ORIGIN
Query Match 19.5%; Score 152; DB 10; Length 391;
Best Local Similarity 100.0%; Pred. No. 7.1e+45;
Matches 152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 507 GAAGATCAGGATACAGCTGAGATCCAGTCGCGCAGCATGGAAGGTGATCTGCAAGAGCTG 566
Db 115 GAAGATCAGGATACAGCTGAGATCCAGTCGCGCAGCATGGAAGGTGATCTGCAAGAGCTG 174

QY 567 CATCAGTCAAAACACCGGGGATAAATCTGGATTGGTTCCGGCCTCAAGGTGAAGATAAT 626
Db 175 CATCAGTCAAAACACCGGGGATAAATCTGGATTGGTTCCGGCCTCAAGGTGAAGATAAT 234

QY 627 ACCTAAGAGGAACACTGTAAATGCCAGAAG 658
```

```
Db 235 ACCTAAGAGGAACACTGTAAATGCCAGAAG 266

RESULT 34
LOCUS AW386180 195 bp mRNA linear EST 04-FEB-2000
DEFINITION RC0-PT0006-271199-011-A08 PT0006 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW386180
VERSION AW386180.1 GI:6890839
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 195)
AUTHORS HCGP http://www.ludwig.org.br/ORESTES.
TITLE The FAPESP/LICR Human Cancer Genome Project
JOURNAL Unpublished (1999)
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC0&t2=RC0-PT0006-
271199-011-A08&t3=1999-11-27&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 195.
FEATURES
source
1..195
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="PT0006"
/dev_stage="Adult"
Note="Organ: pnet; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT 69 a 50 c 56 g 20 t
ORIGIN
Query Match 18.3%; Score 143; DB 9; Length 195;
Best Local Similarity 100.0%; Pred. No. 1.8e+41;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 GACACACAAAACACAGAACACACAGCCAGTCCAGAGCCCGCCAGTAAATGGAGAGCCCA 451
Db 12 GACACACAAAACACAGAACACACAGCCAGTCCAGAGCCCGCCAGTAAATGGAGAGCCCA 71

QY 452 AAAAGAGAACCAGCAGCTGAAAGTCCGGATCCTACACCTGGGCAGCAGACAGAAGA 511
Db 72 AAAAGAGAACCAGCAGCTGAAAGTCCGGATCCTACACCTGGGCAGCAGACAGAAGA 131

QY 512 TCAGGATACAGCTGAGATCCAG 534
Db 132 TCAGGATACAGCTGAGATCCAG 154

RESULT 35
LOCUS AW386185 195 bp mRNA linear EST 04-FEB-2000
DEFINITION RC0-PT0006-271199-011-D10 PT0006 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW386185
VERSION AW386185.1 GI:6890844
KEYWORDS EST.
```

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 195)
AUTHORS HCGP <http://www.ludwig.org.br/ORESTES>.
TITLE The FAPESP/LICR Human Cancer Genome Project
JOURNAL Unpublished (1999)
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL.
(<http://www.ludwig.org.br/scripts/gethtml2.pl?l1=RC0&l2=RC0-PT0006-271199-011-D10&t3=1999-11-27&t4=1>)
Seq primer: puc 18 forward
High quality sequence stop: 195.
Location/Qualifiers
1..195
FEATURES
Source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="PT0006"
/dev_stage="Adult"
/note="Organ: pnet; Vector: puc18; Site1: Small; Site2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 /716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 69 a 50 c 56 g 20 t
ORIGIN
Query Match 18.3%; Score 143; DB 9; Length 195;
Best Local Similarity 100.0%; Pred. No. 1.8e-41;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 392 GACACACAAACACAGACACAGCCAGTCCAGGAGCCAGTAATGGAGGCCCA 451
Db 12 GACACACAAACACAGACACAGCCAGTCCAGGAGCCAGTAATGGAGGCCCA 71
QY 452 AAAAGAAGAACACGACGCTGAAAGTCGGATCCTACACCTGGGCAGCAGACAGAAGA 511
Db 72 AAAAGAAGAACACGACGCTGAAAGTCGGATCCTACACCTGGGCAGCAGACAGAAGA 131
QY 512 TCAGGATACAGCTGACATCCAG 534
Db 132 TCAGGATACAGCTGACATCCAG 154
RESULT 36
BG205090 389 bp mRNA linear EST 21-APR-2001
LOCUS BST24509 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
DEFINITION BG205090
ACCESSION BG205090
VERSION BG205090.1 GI:13726777
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 389)
Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R., Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher,J., Danzig,J. and Ducar,M.
TITLE Creation of genome-wide protein expression libraries using random

JOURNAL activation of gene expression
MEDLINE Nat. Biotechnol. 19 (5), 440-445 (2001)
COMMENT 2127151
Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atersys.com
High quality sequence stop: 389.
Location/Qualifiers
1..389
FEATURES
Source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/cell_line="HT1080"
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."
BASE COUNT 141 a 78 c 97 g 72 t
ORIGIN
Query Match 17.8%; Score 139; DB 10; Length 389;
Best Local Similarity 99.5%; Pred. No. 3.4e-40;
Matches 189; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 535 TCGCGACATGGAAGTGATCTGCAAGAGCTGCATCAACACCGGGGATAATCTG 594
Db 142 TCGCGACATGGAAGTGATCTGCAAGAGCTGCATCAACACCGGGGATAATCTG 201
QY 595 GATTTGGGTTCGGCGCTCAAGGTGAAGATAATACCTAAAGAGGAACACTGTAAAATGCCA 654
Db 202 GATTTGGGTTCGGCGCTCGAGGTGAAGATAATACCTAAAGAGGAACACTGTAAAATGCCA 261
QY 655 GAAGCAGGTGAAGAGCAACCAACCAAGTTTAAATGAAGACAAGCTGAACCAACGCAAGCTGG 714
Db 262 GAAGCAGGTGAAGAGCAACCAACCAAGTTTAAATGAAGACAAGCTGAACCAACGCAAGCTGG 321
QY 715 TTTTATATTA 724
Db 322 TTTTATATTA 331
RESULT 37
T19106
LOCUS D01019T Testis 1 Homo sapiens cDNA clone D01019 5' end, mRNA
DEFINITION T19106
ACCESSION T19106
VERSION T19106.1 GI:601149
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 324)
Pawlak,A., Toussaint,C., Levy,I., Bulle,F., Poyard,M., Barouki,R. and Guellaen,G.
TITLE Characterization of a large population of mRNAs from human testis
JOURNAL Genomics 26, 151-158 (1995)
MEDLINE 95301283
COMMENT Contact: Guellaen G
Unite INSERM 99
INSERM
Unite INSERM 99, Hopital Henri Mondor, 94010 Creteil, France
Tel: (33)149813530
Fax: (33)14980908
Email: guellaen@infobiogen.fr
Seq primer: T7.
Location/Qualifiers

```

source
1. .324
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="p01019"
/clone_lib="Testis 1"
/note="Vector: pGEM 5zf(+); Site_1: EcoRV; Site_2: NotI;
mRNA was prepared from human testis of a 27 years old man.
cDNA was prepared using a 15mer oligo dt anchored by two
degenerated bases at its 3'end and containing a NotI site
at its 5'end. The cDNA was cloned between EcoRV and NotI
sites of pGEM 5zf(+). The 3' end is at the NotI site. The
EcoRV site is lost during the cloning procedure. cDNA
corresponding to abundant species were eliminated from
this library."
BASE COUNT      92 a      81 c      95 g      50 t      6 others
ORIGIN
Query Match      16.0%; Score 125; DB 10; Length 324;
Best Local Similarity 100.0%; Pred. NO. 4.2e-35;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 GACACACACAAACACAGACACAGCCAGTCCCGAGGCCAGTAATGGAGAGCCCA 451
      |||||||
Db 88 GACACACACAAACACAGACACAGCCAGTCCCGAGGCCAGTAATGGAGAGCCCA 147

QY 452 AAAAGAAGAACCCAGCAGTGAAGTCCGATCTACACCTGGCGCAGCAGACAGAAGA 511
      |||||||
Db 148 AAAAGAAGAACCCAGCAGTGAAGTCCGATCTACACCTGGCGCAGCAGACAGAAGA 207

QY 512 TCAGG 516
      |||||
Db 208 TCAGG 212

RESULT 38
AW386181/c
LOCUS      AW386181      132 bp      mRNA      linear      EST 04-FEB-2000
DEFINITION      RC0-PT0006-271199-011-A08_1 PT0006 Homo sapiens cDNA, mRNA
sequence.
ACCESSION      AW386181
VERSION      AW386181.1 GI:6890840
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 132)
HCGP http://www.ludwig.org.br/ORESTES.
The FAPESP/LICR Human Cancer Genome Project
Unpublished (1999)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC0&t2=RC0-PT0006-
271199-011-A08_1&t3=1999-11-27&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 132.
Location/Qualifiers
1. .132
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="PT0006"
/dev_stage="Adult"
/note="Organ: pnet; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196

```

```

,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT      10 a      33 c      38 g      51 t
ORIGIN
Query Match      15.5%; Score 121; DB 9; Length 132;
Best Local Similarity 100.0%; Pred. NO. 1.8e-33;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 GACACACAAACACAGACACAGCCAGTCCCGAGGCCAGTAATGGAGAGCCCA 451
      |||||||
Db 121 GACACACAAACACAGACACAGCCAGTCCCGAGGCCAGTAATGGAGAGCCCA 62

QY 452 AAAAGAAGAACCCAGCAGTGAAGTCCGATCTACACCTGGCGCAGCAGACAGAAGA 511
      |||||||
Db 61 AAAAGAAGAACCCAGCAGTGAAGTCCGATCTACACCTGGCGCAGCAGACAGAAGA 2

QY 512 T 512
      |
Db 1 T 1

RESULT 39
BG208252
LOCUS      BG208252      406 bp      mRNA      linear      EST 21-APR-2001
DEFINITION      R527744 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION      BG208252
VERSION      BG208252.1 GI:13729939
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 406)
Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.,
Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith
,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
,J., Danzig,J. and Ducar,M.
Creation of genome-wide protein expression libraries using random
activation of gene expression
Nat. Biotechnol. 19 (5), 440-445 (2001)
21227151
Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@athersys.com
High quality sequence stop: 406.
Location/Qualifiers
1. .406
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/cell_line="HT1080"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."
BASE COUNT      154 a      83 c      99 g      70 t
ORIGIN
Query Match      15.5%; Score 121; DB 10; Length 406;
Best Local Similarity 100.0%; Pred. NO. 1e-33;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 414 CACAGCCAGTCCCGAGGCCAGTAATGGAGAGCCCAAGACACACAGCTGAA 473
      |||||||

```

Db 22 CACAGCCACTCCAGGAGCCCACTAATCGAGAGCCCAAAAAGAGACCAGCAGCTGAA 81

Qy 474 AGTCGGGATCCTACACCTGGGCAGACAGAGAAGATCAGGATACAGCTGAGATCCCA 533
|||||

Db 82 AGTCGGGATCCTACACCTGGGCAGACAGAGAAGATCAGGATACAGCTGAGATCCCA 141
|||||

Qy 534 G 534

Db 142 G 142

RESULT 40

AW386186/c

LOCUS

DEFINITION AW386186 132 bp mRNA linear EST 04-FEB-2000
RCO-PT0006-271199-011-D10_1 PT0006 Homo sapiens cDNA, mRNA
sequence.

ACCESSION AW386186

VERSION

KEYWORDS AW386186.1 GI:6890845

SOURCE EST.

ORGANISM human.

REFERENCE Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 132)
HCGP <http://www.ludwig.org.br/ORESTES>.
The FAPESP/LICR Human Cancer Genome Project
Unpublished (1999)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL.
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RCO&t2=RCO-PT0006-271199-011-D10.1&t3=1999-11-27&t4=1>)
Seq primer: puc 18 forward
High quality sequence stop: 132.

FEATURES

source

1..132
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="PT0006"
/dev_stage="Adult"
/note="Organ: pnet; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

BASE COUNT 10 a 33 c 38 g 48 t 3 others

ORIGIN

Query Match 14.0%; Score 109; DB 9; Length 132;
Best Local Similarity 100.0%; Pred. No. 3.8e-29;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 392 GACACACAAACACAGAACCCAGCCAGTCCCGAGAGCCAGTAATGGAGAGCCCA 451
|||||

Db 121 GACACACAAACACAGAACCCAGCCAGTCCCGAGAGCCAGTAATGGAGAGCCCA 62
|||||

Qy 452 AAAAGAGAACCAGCAGCTGAAGTCGGGATCCTACACCTGGGCAGCAG 500
|||||

Db 61 AAAAGAGAACCAGCAGCTGAAGTCGGGATCCTACACCTGGGCAGCAG 13
|||||